

Monitoring Nonstationary and Dynamic Trends for Practical Process Fault Diagnosis

¹Yuanling Lin, ²Uwe Kruger, ³Fengshou Gu, ³Andrew Ball, ¹Qian Chen

¹*State Key Laboratory of Mechanics and Control of Mechanical Structures, College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, People's Republic of China*

²*Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180-3590, United States of America*

³*Centre for Efficiency and Performance Engineering, University of Huddersfield, Huddersfield, HD1 3DH, United Kingdom*

Abstract

This article introduces a revised common trend framework to monitor nonstationary and dynamic trends in industrial processes and shows needs for each improvement on the basis of three application studies. These improvements relate to (i) the extension of the common trend framework to include sets that contain stationary and nonstationary variables, (ii) handling cases where residuals are not drawn from multivariate normal distributions and (iii) the application of the framework to larger variable sets. Existing work does not adequately address these practically important issues. Industrial application studies highlight the needs for (i) the extended framework to model data sets containing stationary and nonstationary variables, (ii) handling statistics that are not based on normally distributed residuals and (iii) the use of Chigira procedure to robustly extract common trends. The extended framework is compared to traditional approaches.

Keywords

common trends models; Kasa decomposition; cointegration; latent stationary and nonstationary factors; generalized data structure; forecast recovery; multivariate normal distribution

1 Introduction

The development of instrumentation and automation for modern industrial processes in the chemical and general manufacturing industries allows large quantities of data to be utilized for assessing current operating conditions (Kruger & Xie, 2012; Ge & Song, 2013; Severson, Chaiwatanodom & Braatz, 2016). Traditional approaches to monitor general processes include model-based (Ding, 2013; Zhong, Xue & Ding, 2018; Liu, Luo, Yang & Wu, 2016; Li, Gao, Shi & Lam, 2016; Zhao, Yang, Ding & Li, 2018), signal-based (Lei, Lin, He & Zuo, 2013; Yan, Gao & Chen, 2014; Fan, Cai, Zhu, Shen, Huang & Shang, 2015; Wu, Guo, Xie, Ni & Liu, 2018), and knowledge-based (Gao, Cecati & Ding, 2015; Mohammadi & Montazeri-Gh, 2015; Chiremsel,

* Corresponding Authors: +86-25-8489-3221, q.chen@nuaa.edu.cn (Qian Chen); +1-518-276-4818, krugeu@rpi.edu (Uwe Kruger)

Said & Chiremsel, 2016; Davies, Jackson & Dunnett, 2017) techniques. Based on their conceptual simplicity, techniques that relate to multivariate statistical process control (MSPC) (Kruger & Xie, 2012; Ge & Song, 2013; Qin, 2012; Ge, Song & Gao, 2013; Yin, Li, Gao & Kaynak, 2015) have also gained attention over the past few decades, particularly for applications to industrial processes that produce larger variable sets.

At present, MSPC-based techniques are firmly based on core component technologies that predominantly include principal component analysis, partial least squares, independent component analysis and their extensions (Rato, Reis, Schmitt, Hubert & De Ketelaere, 2016; Li, Qin & Zhou, 2010; Lee, Qin & Lee, 2006; Yu, Khan & Garaniya, 2016; Zhang, Sun & Fan, 2015; Li & Yang, 2015; Cai, Tian & Chen, 2017). These techniques, however, rely on the fundamental assumptions that (i) the process operates at a predefined stationary operating condition, where the recorded variables set has a joint multivariate distribution with a time-invariant mean vector and covariance matrix, and (ii) the recorded set possess no serial correlation. However, such assumptions are rarely satisfied in industrial practice, as recorded process variables usually describe (i) time-varying behaviors that involve changes in operating conditions, variations in process feeds, emptying and filling cycles, the presence of unmeasured disturbances, and operator interventions, and (ii) controller feedback which produces inherently correlated process variables.

To address serial correlation among and between measured process variables, the literature proposed a number of dynamic extensions for standard MSPC methods (Ku, Storer, & Georgakis, 1995; Li, Qin, & Zhou, 2014; Dong & Qin, 2018; Lee, Yoo, & Lee, 2004; Huang & Yan, 2015; Fan & Wang, 2014). Conversely, the problem of nonstationary process variables has only been sporadically addressed. Adapting slowly time-varying means and variances of process variables can accommodate mildly nonstationary behavior of process units (Wang, Kruger, & Lennox, 2003; Wang, Kruger, & Irwin, 2005; Xie, Li, Zeng, & Kruger, 2016; Jaffel, Taouali, Harkat, & Messaoud, 2016). This, however, may accommodate incipient fault conditions, which is undesirable. Another, more promising avenue is employing cointegration-based models, where the stationary cointegration relationships among nonstationary process variables are utilized for process monitoring (Chen, Kruger & Leung, 2009; Sun, Zhang, Zhao & Gao, 2017). However, the use of long-term equilibria, *i.e.*, the cointegration relationships, is insensitive in detecting anomalous events that are orthogonal to the cointegration space.

More recently, a generic cointegration framework has been proposed, which can handle nonstationary variables that are serially correlated (Lin, Kruger & Chen, 2017). This technique (i) separates nonstationary and stationary components from the nonstationary variable set on the basis of a cointegration model (Engle & Granger, 1987; Pfaff, 2008), (ii) extracts both latent nonstationary and stationary factors using a common trends model (Stock & Watson, 1988; Kasa, 1992), and (iii) defines two statistics for monitoring the two types of factors individually. Although the detection of abnormal behavior in industrial nonstationary dynamic

processes has been presented through the application to a glass melter process (Lin, Kruger & Chen, 2017), the following practically relevant issues have not been considered:

- the application studies presented herein contain recorded variable sets that are both, nonstationary and stationary, whereas Lin, Kruger & Chen (2017) addressed the problem of variable sets that are nonstationary only;
- the model residuals may not be drawn from multivariate normal distributions, in contrast to the work in Lin, Kruger & Chen (2017) where this assumption is made; and
- the conventional cointegration testing procedure by Johansen (1995) may run into difficulties if the number of variables is large, which is typically the case for complex industrial processes.

This article is divided into the following sections. Section 2 gives a detailed description of the processes, including the FCCU simulator, the polymerization and the distillation process, used in this article. Section 3 then summarizes the monitoring framework and introduces an improved common trends framework. Sections 4 to 6 summarize the applications to the three processes. A concluding summary is given in Section 7.

2 Process Descriptions

This section provides a detailed description of the three processes that are used to demonstrate the working of the improved cointegration framework. Subsection 2.1 gives a description of the simulated fluid catalytic cracking unit. Subsections 2.2 and 2.3 then provide details of the polymerization and distillation processes, respectively.

2.1 Fluid Catalytic Cracking Unit

A fluid catalytic cracking unit (FCCU) is an important part of a refinery. An FCCU receives several different heavy feed stocks from other refinery units and cracks these into more valuable components that are usually blended into gasoline and other products. The main feed stream to an FCCU is gas oil, but heavier diesel and wash oil streams also contribute to the total feed stream that is first preheated in a heat exchanger and furnace and then passed to the riser. The total feed stream, mixed with slurry from the main fractionator bottoms, is mixed with hot, regenerated catalyst from the regenerator in the riser. The hot catalyst provides the necessary heat for the endothermic reactions producing gaseous reactants that pass through the main fractionator for separation. Resulting from the cracking reactions, a carbonaceous material, coke, deposits on the surface of the catalyst and depletes its catalytic property. Recycling the spent catalyst to the regenerator, where it is mixed with air in a fluidized bed, regenerates the catalyst. Air, pumped to the regenerator with a high capacity combustion air blower and a smaller lift air blower, reacts with coke and produces carbon monoxide and

carbon dioxide. McFarlane, Reineman, Bartee & Georgakis (1993) provide a very detailed discussion of the FCCU simulator. Table 1 lists the 28 variables that were analyzed in Section 4.

[Table 1 about here]

2.2 Polymerization Process

This process unit is a spinning cell that is part of a large polymerization process. Polymers of synthetic fibers, for example nylon and polyethylene terephthalate, are usually produced by complex chemical reactions based on petroleum and natural gas. Before the raw product enters the spinning cell, it is molten and pumped through a spinneret, which the fluid polymer exits through a set of holes in the form of filaments. The filaments are then cooled by air, which converts them into a rubbery state first and finally into a solid state. The spinning process involves extrusion and the subsequent solidification, after which the filaments are combined to produce threads that are collected on a take-up wheel. The fibers, when stretched in both the molten and solid states, define the orientation of the polymer chains along the fiber axis. A more detailed discussion concerning the operation of spinning cells may be available in Kricheldorf, Nuyken & Swift (2005). Table 2 lists the 7 variables that were recorded for the spinning cell analyzed in Section 5.

[Table 2 about here]

2.3 Distillation Unit

This unit purifies butane from a mixture of heavier hydrocarbons, mainly butane, pentane and impurities of propane that enter the unit as fresh feed. The separation is achieved by trays in the distillation column. Butane is taken off the top of the column and the heavier hydrocarbons are drawn from the bottom of the column. Large fan condensers condense the overhead stream and the liquid petroleum gas is fed to a reflux drum. The reflux ratio is set to meet predefined impurity levels in the product stream. The bottom stream contains heavier hydrocarbons and is split between a stream that is fed back to the column through a reboiler and a feed to a downstream processing operation. A sufficient liquid level in the reboiler must be maintained to ensure that the steam coils are immersed to avoid accretion of the coils. Table 3 lists the 12 variables that were analyzed in Section 6.

[Table 3 about here]

3 Modeling and Monitoring Nonstationary and Dynamic Trends

This section introduces the common trends framework and presents the practical refinements required to apply it in practice. After briefly describing the underlying data structure in Subsection 3.1, Subsection 3.2 summarizes the cointegration theory and then demonstrates how the common trends model can be applied in separating nonstationary components into latent nonstationary and stationary factors. This is followed by highlighting how to construct monitoring statistics and discussing the issue of forecast recovery in Subsections 3.3 and 3.4, respectively. Finally, Subsection 3.5 outlines the practically important improvements for applying the common trends framework, which constitutes the main contributions in this article.

3.1 Data Structure

Let $\mathbf{y} = (y_1 \ y_2 \ \cdots \ y_n)^T$ be a vector storing a recorded variable set of dimension n , generated by a set of non-observed factors. We assume that each component in \mathbf{y} consists of linear combinations of a set of serially correlated nonstationary and stationary factors. Defining z as the time-shift operator, the corresponding data model for \mathbf{y} is as follows:

$$\mathbf{y}(z) = \mathbf{A}\mathbf{t}(z) + \mathbf{B}\mathbf{u}(z) = [\mathbf{A} \ \mathbf{B}] \begin{pmatrix} \mathbf{t}(z) \\ \mathbf{u}(z) \end{pmatrix}, \quad (1)$$

where $\mathbf{t} \in \mathbb{R}^{n_{ns}}$ represents a set of nonstationary factors, $\mathbf{u} \in \mathbb{R}^{n_s}$ is a set of stationary factors, and $\mathbf{A} \in \mathbb{R}^{n \times n_{ns}}$ and $\mathbf{B} \in \mathbb{R}^{n \times n_s}$ are parameter matrices that define the contributions of \mathbf{t} and \mathbf{u} to the recorded process variables \mathbf{y} , respectively. Without restriction of generality, we assume that $n_{ns} + n_s \leq n$. The random vectors \mathbf{t} and \mathbf{u} are modeled by a vector auto-regressive integrated moving average, VARIMA, model and a vector auto-regressive moving average, VARMA, model, respectively:

$$\mathbf{C}_{ns}(z)\nabla^d \mathbf{t}(z) = \mathbf{D}_{ns}(z)\mathbf{e}_1(z) \quad \mathbf{C}_s(z)\mathbf{u}(z) = \mathbf{D}_s(z)\mathbf{e}_2(z), \quad (2)$$

where $\mathbf{C}_{ns}(z) = \mathbf{I} + \mathbf{C}_{ns,1}z^{-1} + \mathbf{C}_{ns,2}z^{-2} + \cdots + \mathbf{C}_{ns,p_{ns}}z^{-p_{ns}}$ and $\mathbf{D}_{ns}(z) = \mathbf{I} + \mathbf{D}_{ns,1}z^{-1} + \mathbf{D}_{ns,2}z^{-2} + \cdots + \mathbf{D}_{ns,q_{ns}}z^{-q_{ns}}$ are matrix polynomial functions of dimension $n_{ns} \times n_{ns}$, $\mathbf{C}_s(z) = \mathbf{I} + \mathbf{C}_{s,1}z^{-1} + \mathbf{C}_{s,2}z^{-2} + \cdots + \mathbf{C}_{s,p_s}z^{-p_s}$ and $\mathbf{D}_s(z) = \mathbf{I} + \mathbf{D}_{s,1}z^{-1} + \mathbf{D}_{s,2}z^{-2} + \cdots + \mathbf{D}_{s,q_s}z^{-q_s}$ are matrix polynomial functions of dimension $n_s \times n_s$, $\nabla = \mathbf{I}(1 - z^{-1})$, d is the integration order of the nonstationary factors, and the stationary and serially uncorrelated sequence \mathbf{e}_i ($i = 1, 2$) are assumed to be drawn from multivariate normal distributions with zero means and full rank covariance matrices \mathbf{R}_{e_i} , i.e., $\mathbf{e}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{e_i})$.

3.2 Cointegration and Common trends

Engle & Granger (1987) coined the term cointegration to describe that a nonstationary variable set can have linear combinations that are stationary. Let $\boldsymbol{\beta}_i \in \mathbb{R}^n$ be a parameter vector that yields:

$$\epsilon_i = \boldsymbol{\beta}_i^T \mathbf{y}, \quad (3)$$

where ϵ_i is a serially correlated stationary sequence that is referred to as a cointegration residual. It should be noted that there can be as many as $n - 1$ different cointegrating vectors. If there are r ($1 \leq r \leq n - 1$) such linearly independent vectors $\boldsymbol{\beta}_i$ ($i = 1, 2, \dots, r$), \mathbf{y} is said to be cointegrated with a cointegrating rank of r , or there are r cointegration relationships embedded within the variable set \mathbf{y} . This allows defining the cointegrating matrix $\boldsymbol{\beta} = [\boldsymbol{\beta}_1 \quad \boldsymbol{\beta}_2 \quad \dots \quad \boldsymbol{\beta}_r] \in \mathbb{R}^{n \times r}$.

Econometricians and statisticians have developed a substantial body of research in developing the theory of cointegration and presented numerous applications. This framework includes estimating the cointegrating rank and the cointegrating vectors and for applications to pairs of nonstationary variables, the Engle-Granger two-step procedure (Engle & Granger, 1987) has been proposed. For $n > 2$, the Johansen procedure (Johansen, 1995) has been extensively used for constructing cointegration models.

Given that cointegrated variables share common trends, Stock & Watson (1988) derived a representation for reformulating the cointegration model as a common trends model (Stock & Watson, 1988; Kasa, 1992; Gonzalo & Granger, 1995; Escribano & Peña, 1994). Given that \mathbf{y} is a vector of n nonstationary variables that have a cointegrating rank r , the common trends representation of \mathbf{y} is a linear combinations of $n - r$ nonstationary trends, *i.e.*, common trends or common factors, and a set of stationary components:

$$\mathbf{y} = \mathbf{y}_{ns} + \mathbf{y}_s = \boldsymbol{\beta}_\perp \mathbf{f} + \mathbf{y}_s, \quad (4)$$

where $\boldsymbol{\beta}_\perp \in \mathbb{R}^{n \times (n-r)}$ represents the orthogonal complement of the cointegrating matrix $\boldsymbol{\beta}$, $\mathbf{f} \in \mathbb{R}^{n-r}$ is a set of nonstationary variables that have no cointegration relationships and describe the common trends, and \mathbf{y}_s is the stationary components of \mathbf{y} . Moreover, Escribano & Peña (1994) proved that \mathbf{y} is driven by $n - r$ nonstationary factors and r stationary factors, which implies that Eq. (4) is equivalent to a common trends model in the form of the Kasa decomposition (Kasa, 1992):

$$\mathbf{y}(k) = \boldsymbol{\beta}_\perp [\boldsymbol{\beta}_\perp^T \boldsymbol{\beta}_\perp]^{-1} \boldsymbol{\beta}_\perp^T \mathbf{y}(k) + \boldsymbol{\beta} [\boldsymbol{\beta}^T \boldsymbol{\beta}]^{-1} \boldsymbol{\beta}^T \mathbf{y}(k), \quad (5)$$

where $\beta_{\perp}^T \mathbf{y}$ and $\beta^T \mathbf{y}$ are the extracted nonstationary and stationary factors, respectively. For the model in Eq. (5), β_{\perp} is referred to as the common trends loading matrix, and the identified nonstationary factors $\beta_{\perp}^T \mathbf{y}$ constitute the non-cointegrated part of \mathbf{y} .

3.3 Monitoring Statistics

The decomposition described in Subsection 3.2 allows identifying the nonstationary factors $\mathbf{f} = \beta_{\perp}^T \mathbf{y}$ and the stationary factors $\mathbf{g} = \beta^T \mathbf{y}$ from the recorded variable set \mathbf{y} . Comparing Eqs. (1) and (5), it follows that $\mathbf{f} \propto \mathbf{t}$ and $\mathbf{g} \propto \mathbf{u}$, implying that the vector spaces spanned by \mathbf{f} and \mathbf{g} converge, in probability, to the vector spaces spanned by \mathbf{t} and \mathbf{u} as the sample size of the recorded variable set approaches infinity, respectively – in terms of the statistical properties of the first and second moments (Chan & Wei, 1988; Tanaka, 1996). Hence, the terms on the right hand side of Eq. (5) converge in probability to $\mathbf{A}\mathbf{t}$ and $\mathbf{B}\mathbf{u}$. Following from Eq. (1), both the nonstationary factors \mathbf{t} and stationary factors \mathbf{u} are serially correlated. Kruger, Zhou & Irwin (2004) and Xie, Kruger, Lieftucht, Littler, Chen & Wang (2006), however, demonstrated that serial correlation among and between the random variables can render nonnegative quadratic test statistics to be either insensitive or to produce excess numbers of false alarms. This is because the assumption that the recorded data points are drawn independently, *i.e.*, the k th data point is independently drawn from the $(k - 1)$ th data point, is violated.

To address this issue, Kruger, Zhou & Irwin (2004) and Xie, Kruger, Lieftucht, Littler, Chen & Wang (2006) proposed identifying a vector auto-regressive moving average, VARMA, model and construct nonnegative test statistics on the basis of the model residuals, which are not serially correlated. Moreover, the Wold decomposition theorem (Box, Jenkins, Reinsel & Ljung, 2016) proves that a VARMA model can be converted into a vector auto-regressive (VAR) model. In a similar fashion, it is possible to convert a vector auto-regressive integrated moving average (VARIMA) model to an equivalent vector auto-regressive integrated (VARI) model. VARI and VAR models can describe serially correlated nonstationary and stationary signals, respectively. Therefore, the random vectors \mathbf{f} and \mathbf{g} can be described by VARI and VAR model structures, which can be identified using the computationally efficient Yule-Walker equation.

The identified nonstationary factors \mathbf{f} are assumed to be integrated of order $d = 1$. Defining k as a time-based index, Appendix A shows that a VARI model for the nonstationary factors is given by:

$$\mathbf{f}(k) = \sum_{j=1}^{p_{\text{ns}}+1} \Psi_j \mathbf{f}(k-j) + \mathbf{e}_f(k), \quad (6)$$

where p_{ns} is the number of lagged terms, $\Psi_j, j = 1, \dots, p_{\text{ns}}, p_{\text{ns}} + 1$ are coefficient matrices, and \mathbf{e}_f denotes a random error vector that has a multivariate normal distributions with a mean of zero and the covariance

matrix \mathbf{R}_f . As before, the observations of the random vector \mathbf{e}_f are independent and identically distributed, and do not contain any serial correlation. Hence, the following Hotelling's T^2 statistic (Johnson & Wichern, 2007) can be constructed to monitor the nonstationary factors in \mathbf{y} :

$$T_{ns}^2 = \mathbf{e}_f^T \hat{\mathbf{R}}_f^{-1} \mathbf{e}_f, \quad (7)$$

Here, $\hat{\mathbf{R}}_f$ is the sample covariance matrix of \mathbf{e}_f , estimated from a total of N reference points, *i.e.*, $\hat{\mathbf{R}}_f = N^{-1} \sum_{k=1}^N \mathbf{e}_f(k) \mathbf{e}_f^T(k)$. If the random vector \mathbf{e}_f has a multivariate normal distribution, the critical value, or control limit, for the test statistic T_{ns}^2 to reject the null hypothesis H_0 : the process is in-statistical-control, for a significance α , is given by a scaled F -distribution (Kruger & Xie, 2012):

$$T_{ns,\alpha}^2 = \frac{n_{ns}(N-1)}{N-n_{ns}} F_{\alpha}(n_{ns}, N - n_{ns}), \quad (8)$$

where $F_{\alpha}(n_{ns}, N - n_{ns})$ is critical value of an F -distribution with n_{ns} and $N - n_{ns}$ degrees of freedom.

Similarly, the VAR model structure for the identified stationary factors \mathbf{g} is as follows:

$$\mathbf{g}(k) = \sum_{j=1}^{p_s} \Phi_j \mathbf{g}(k-j) + \mathbf{e}_g(k), \quad (9)$$

where p_s is the number of lagged terms, $\Phi_j, j = 1, \dots, p_s$, are coefficient matrices and \mathbf{e}_g denotes a random error vector that has a multivariate normal distributions with a mean of zero and the covariance matrix \mathbf{R}_g . It is important to note that the random vector \mathbf{e}_g does not possess any serial correlation, implying that realizations of \mathbf{e}_g are independent and identically distributed. Thus, the Hotelling's T^2 statistic (Johnson & Wichern, 2007) is constructed for monitoring the variation of the stationary components of \mathbf{y} :

$$T_s^2 = \mathbf{e}_g^T \hat{\mathbf{R}}_g^{-1} \mathbf{e}_g. \quad (10)$$

Here, $\hat{\mathbf{R}}_g$ is the sample covariance of the random vector \mathbf{e}_g , estimated from a total of N reference points, *i.e.*, $\hat{\mathbf{R}}_g = N^{-1} \sum_{k=1}^N \mathbf{e}_g(k) \mathbf{e}_g^T(k)$. If the random vector \mathbf{e}_g has a normal distribution, the critical value, or control limit, for testing H_0 : the process is in-statistical-control, using the test statistic in Eq. (10), for a significance of α , is given by:

$$T_{s,\alpha}^2 = \frac{n_s(N-1)}{N-n_s} F_\alpha(n_s, N - n_s), \quad (11)$$

where $F_\alpha(n_s, N - n_s)$ is the critical value of an F -distribution with n_s and $N - n_s$ degrees of freedom. The alternative hypotheses for the test statistics in Eqs. (7) and (10) are H_1 : the process is-out-of-statistical-control.

3.4 Forecast Recovery and Fault Magnitude Compensation

A practically important issue that relates to the diagnosis of anomalous events using model residuals is forecast recovery. More precisely, the removal of correlation by applying vector auto-regressive models generates the problem of forecast recovery (Apley & Shi, 1999; Superville & Adams, 1994). Considering the VARI model in Eq. (6) and computing the difference between the k th data point and its prediction yields:

$$\mathbf{e}_f(k) = \mathbf{f}(k) - \sum_{j=1}^{p_{ns}+1} \boldsymbol{\Psi}_j \mathbf{f}(k-j). \quad (12)$$

In the presence of a fault condition, assuming a step-type fault of magnitude $\Delta \mathbf{f}$ (for simplicity) that arises at the k th time index, we have:

$$\mathbf{e}_f^{(f)}(k) = \mathbf{e}_f(k) + \Delta \mathbf{f} = [\mathbf{f}(k) + \Delta \mathbf{f}] - \sum_{j=1}^{p_{ns}+1} \boldsymbol{\Psi}_j \mathbf{f}(k-j), \quad (13)$$

Given that the diagonal matrix ∇ contains the backward finite difference term $1 - z^{-1}$, *i.e.*, $\nabla \mathbf{f}^{(f)}(k+m) = \nabla[\mathbf{f}(k+m) + \Delta \mathbf{f}] = \nabla \mathbf{f}(k+m)$, it follows that $\lim_{m \rightarrow +\infty} [\mathbf{e}_f^{(f)}(k+m) - \mathbf{e}_f(k+m)] = \mathbf{0}$. This phenomena is referred to as forecast recovery and implies that differencing a step-type fault does not manifest itself in a step-type fault signature of the residual sequence $\mathbf{e}_f^{(f)}(k+m)$. This can be circumvented by utilizing the following filter mechanism (Lieftucht, Kruger, Xie, Littler, Chen & Wang, 2006):

$$\bar{\mathbf{e}}_f^{(f)}(k) = \sum_{j=0}^k \mathbf{e}_f^{(f)}(j), \quad \bar{\mathbf{f}}(k) = \mathbf{f}^{(f)}(k) - \bar{\mathbf{e}}_f^{(f)}(k). \quad (14)$$

More precisely, instead of using the realization $\mathbf{f}^{(f)}(k)$, the filtered realization $\bar{\mathbf{f}}(k)$ is used. This compensation scheme is utilized in this article for estimating the correct magnitudes and signatures of general deterministic fault conditions.

3.5 Practically Important Improvements of Common Trends Framework

This subsection discusses three important issues that arise when applying the common trends framework: (i) how to address cases where the variable set contains nonstationary as well as stationary variables, (ii) how to handle model residuals that are not drawn from a multivariate normal distribution and (iii) how to obtain the cointegrating rank, the cointegration matrix and its orthogonal complement for larger variable sets.

Generalized Data Structure and Common Trends Model

The common trends model assumes that all random variables stored in \mathbf{y} are nonstationary. However, a recorded variable set for an industrial process system may contain both nonstationary and stationary variables. As an example, a controlled variable is expected to be stationary, whilst a feedstream may be nonstationary. Whether a random variable is nonstationary or not can be tested using the augmented Dickey-Fuller or ADF test, which is briefly summarized in Appendix B.

To be practically useful, the assumption that all variables in \mathbf{y} are nonstationary need to be relaxed, leading to a generalized data structure. It should be noted that the i th ($1 \leq i \leq n$) component of the recorded variable set y_i in the term of Eq. (1) can be either nonstationary or stationary, whilst the data structure in Lin, et al. (2017) only assumed that all the components of \mathbf{y} have to be nonstationary with the same integration order. If the variable y_i is stationary, it means that latent nonstationary factors \mathbf{t} in the term of Eq. (1) have no impact on y_i and the elements of the i th row vector of the nonstationary factor loading matrix \mathbf{A} in the term of Eq. (1) have to be all zeros, because the sum of a nonstationary and a stationary variable leads to nonstationarity. This yields a two-step procedure for identifying nonstationary and stationary factors from the recorded set that can contain both nonstationary and stationary variables. First, the recorded variable set \mathbf{y} is divided into $\mathbf{y}_{\text{ns}}^{(g)}$ and $\mathbf{y}_s^{(g)}$ which contain the nonstationary and stationary subsets of \mathbf{y} , respectively.

Next, similar to Eq. (5), applying the Kasa decomposition allows extracting nonstationary and stationary factors from $\mathbf{y}_{\text{ns}}^{(g)}$ by determining factor loading matrices $\boldsymbol{\beta}^{(g)}$ and $\boldsymbol{\beta}_\perp^{(g)}$. Then, the stationary components associated with $\mathbf{y}_{\text{ns}}^{(g)}$, i.e., $\left[\mathbf{I} - \boldsymbol{\beta}_\perp^{(g)} \left[\boldsymbol{\beta}_\perp^{(g)\text{T}} \boldsymbol{\beta}_\perp^{(g)} \right]^{-1} \boldsymbol{\beta}_\perp^{(g)\text{T}}\right] \mathbf{y}_{\text{ns}}^{(g)} = \boldsymbol{\beta}^{(g)} \left[\boldsymbol{\beta}^{(g)\text{T}} \boldsymbol{\beta}^{(g)} \right]^{-1} \boldsymbol{\beta}^{(g)\text{T}} \mathbf{y}_{\text{ns}}^{(g)}$ can be combined with $\mathbf{y}_s^{(g)}$ to subsequently extract the stationary factors by estimating the generalized factor loading matrix $\boldsymbol{\beta}$ and its orthogonal complement, i.e., $\boldsymbol{\beta}_\perp = \left(\boldsymbol{\beta}_\perp^{(g)\text{T}} \quad \mathbf{0}^{\text{T}} \right)^{\text{T}}$. With respect to Eqs. (1) and (5), the generalized data structure, referred to by the superscript (g), and its corresponding Kasa decomposition are:

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_{\text{ns}}^{(g)} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{y}_s^{(g)} \end{pmatrix} \Rightarrow \mathbf{y} = \begin{pmatrix} \boldsymbol{\beta}_\perp^{(g)} \\ \mathbf{0} \end{pmatrix} \left[\begin{pmatrix} \boldsymbol{\beta}_\perp^{(g)} \\ \mathbf{0} \end{pmatrix}^{\text{T}} \begin{pmatrix} \boldsymbol{\beta}_\perp^{(g)} \\ \mathbf{0} \end{pmatrix} \right]^{-1} \begin{pmatrix} \boldsymbol{\beta}_\perp^{(g)} \\ \mathbf{0} \end{pmatrix}^{\text{T}} \begin{pmatrix} \mathbf{y}_{\text{ns}}^{(g)} \\ \mathbf{y}_s^{(g)} \end{pmatrix} + \boldsymbol{\beta} \left[\boldsymbol{\beta}^{\text{T}} \boldsymbol{\beta} \right]^{-1} \boldsymbol{\beta}^{\text{T}} \begin{pmatrix} \mathbf{y}_{\text{ns}}^{(g)} \\ \mathbf{y}_s^{(g)} \end{pmatrix} \quad (15)$$

Each of the applications studies, detailed in Sections 4 to 6, exemplify that the generalized data structure and its corresponding Kasa decomposition is essential for the successful application of the common trends framework.

Chigira Procedure

The conventional cointegration testing method for $n > 2$ is the Johansen procedure, which has been widely used for modeling nonstationary processes. However, two important practical issues render the Johansen procedure difficult to apply for constructing the model in Eq. (5). The first issue relates to the fact that critical values of most cointegration tests depends on the dimension of the recorded set, n , and the cointegration rank under the null hypothesis, r . The critical values, obtained empirically and reported in MacKinnon (1991), are not available for large n . The second issue is the complexity of the cointegration models. More precisely, the Johansen procedure depends on auto-regressive time series models and, hence, there may be hundreds of parameters to be estimated if the variable number is large.

To directly address these issues, Chigira (2008) proposed a testing procedure for estimating the cointegration rank on the basis of principal component analysis. This testing procedure relies on the fact that the smallest r principal component loading vectors converge, asymptotically, to the cointegrating vectors whilst the largest $n - r$ principal component loading vectors converge, asymptotically, to orthogonal complements of cointegrating vectors.

Let \mathbf{C}_y be the sample covariance matrix of the n -dimensional observed nonstationary time series $\mathbf{y}(k)$ and \mathbf{B} be principal component loading vectors such that $\mathbf{C}_y = \mathbf{B}\mathbf{A}\mathbf{B}^T$, where \mathbf{A} is a diagonal matrix that consists of eigenvalues. \mathbf{B} can be divided into two component matrices $\mathbf{B} = [\mathbf{B}_{n-r} \quad \mathbf{B}_r]$, where \mathbf{B}_{n-r} and \mathbf{B}_r are the principal component loading vectors corresponding to the largest $n - r$ and the smallest r eigenvalues, respectively. According to the convergence proof in Chigira (2008), $\mathbf{B}_{n-r} \xrightarrow{p} \boldsymbol{\beta}_\perp \mathbf{Q}$ and $\mathbf{B}_r \xrightarrow{p} \boldsymbol{\beta} \mathbf{R}$ as the sample size goes infinity, where \mathbf{Q} and \mathbf{R} are $(n - r) \times (n - r)$ and $r \times r$ constant matrices, respectively. A more detailed convergence proof may be found in Chigira (2008).

Put differently, if the sample size of the reference data set is large enough, \mathbf{B}_r is an estimate of $\boldsymbol{\beta}$ up to a similarity transformation, such that $\mathbf{B}_{n-r}^T \mathbf{y}(k)$ is nonstationary whilst $\mathbf{B}_r^T \mathbf{y}(k)$ is stationary. Thus, the problem of estimating the cointegration rank becomes a problem of verifying the stationarity of principal components. On the basis of these properties, Chigira (2008) developed a sequential procedure to test the cointegration rank by applying a unit root test (Dickey & Fuller, 1981; Phillips & Perron, 1988; Fan & Gencay, 2010)

to verifying whether or not a principal component is nonstationary. The null hypothesis and the alternative are:

H_0 : the cointegration rank is r ; and

H_1 : the cointegration rank is $r - 1$.

which can be reformulated as follows:

H_0 : the $(n - r + 1)$ th principal component is stationary;

H_1 : the $(n - r + 1)$ th principal component is nonstationary.

In practice, the testing procedure commences by testing $r = n$, that is, apply a unit root test to verify the stationarity of the first principal component. If the null hypothesis is accepted, the cointegration rank is determined to be $r = n$. If it is rejected, the hypothesis becomes $r = n - 1$, or the stationarity of the second principal component is tested. The procedure carries out sequentially until a null hypothesis $r = n - r + 1$ is accepted. In the testing procedure, a unit root test is applied to verify the stationarity of the $(n - r + 1)$ th principal component. More details of the procedure implement can be found in Chigira (2008).

In contrast to the commonly used Johansen's framework, the Chigira procedure (i) is not compromised by larger numbers of process variables, whilst the critical values of the Johansen procedure are unavailable for larger n and (ii) unlike the Johansen procedure, the Chigira procedure does not require a pre-identified autoregressive model for each recorded process variable. Moreover, the Monte Carlo experiments reported in Chigira (2008) show that the proposed Chigira procedure performs well for large cointegrated systems even if the sample size is small. Hence, we utilized the Chigira procedure for constructing the model structure in Eq. (15) for processes with a larger number of recorded variables, an example of which is given in Section 4. Conversely, both the Johansen and Chigira procedures showed to work well for smaller variable sets, which is exemplified in Sections 5 and 6.

Chigira (2008) confirmed by using Monte Carlo simulations that the Chigira procedure performs better in terms of correctly estimating the cointegrating rank for small sample sizes. For larger sample sizes, however, both procedures produced identical estimates. Following from the preceding discussion, the Chigira procedure is, in our view, conceptually simpler, robust and closely related to the traditional MSPC framework. We therefore advocate for the use of the Chigira procedure in industrial practice, as (i) it can deal with larger numbers of variables (exceeding 10) and (ii) and it is conceptually simple and robust.

Control Limits for Monitoring Statistics

The two monitoring statistics, T_{ns}^2 and T_s^2 , defined in Subsection 3.3, are F distributed, as they are nonnegative quadratic statistics based on filtered sequences that are assumed to be drawn from a multivariate normal distribution. In industrial practice, however, the filtered residual sequences may not be drawn from a multivariate normal distribution, which can be tested by the well known Anderson-Darling (Anderson & Darling, 1954), Shapiro-Wilk (Shapiro & Wilk, 1965) or Jarque-Bera (Jarque & Bera, 1987) tests. If a sample is not drawn from a normally distributed population, the control limits of the nonnegative squared statistics in Eqs. (8) and (11) are incorrect. To address this problem, the control limits of the T_{ns}^2 and T_s^2 statistics can be obtained from the probability density function that is estimated using kernel density estimation (KDE) (Silverman, 1986; Chen, Wynne, Goulding & Sandoz, 2000). A univariate estimator for a kernel function $K(\cdot)$ is defined by:

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^N K\left\{\frac{x-x_i}{h}\right\}. \quad (16)$$

Here, x is the data point under consideration, x_i is a recorded reference point, h is the smoothing parameter, N is the sample size, and K is the kernel function. The kernel function K must satisfy:

$$\int_{-\infty}^{\infty} K(x)dx = 1. \quad (17)$$

In practice, the Gaussian kernel function, adopted in this article, is widely used.

From the estimated PDF, the control limits can be calculated as follows. First, the T_{ns}^2 and T_s^2 values are computed from the reference data set. Then the univariate kernel density estimator is used to estimate the density functions of the T_{ns}^2 and T_s^2 statistics. Finally, the area under the PDFs that is equal to 0.99 can be determined, and the upper boundaries of the numerical integration become the control limit of the T_{ns}^2 and T_s^2 statistics. Section 6 discusses an application to an industrial data set that does not yield residuals sequences that are drawn from a multivariate normal distribution and demonstrates that an incorrect control limit can have a profound effect on the process monitoring application.

Figure 1 presents a flowcharts summarizing the monitoring approach, which directly addresses practical issues that existing work in this area does not.

[Figure 1 about here]

4 Application to a Fluid Catalytic Cracking Unit

This section summarizes the application of the improved monitoring scheme to the FCCU simulator, described in Subsection 2.1. Subsection 4.1 gives details to guarantee that the simulated data set reflects a practically realistic scenario. Subsection 4.2 shows how to establish monitoring models. Subsection 4.3 then shows how to detect two abnormal events. Finally, Subsection 4.4 discusses the practical importance of the improvements detailed in Subsection 3.5.

4.1 Details of Simulation

To simulate a practically realistic behavior, some of the heavy feeds, *i.e.*, the slurry and wash oil feed, the temperature of the total feed after leaving the preheater and the airflow of the combustion air blower were simulated to be mildly nonstationary. In addition to that, the riser temperature was controlled by the PI controller detailed in Kruger & Xie (2012, Eq. 7.49). McFarlane, Reineman, Bartee & Georgakis (1993) describe other regulatory controllers, for example to guarantee a smooth catalyst flow between the riser and regenerator. To guarantee that the simulator describes a practically relevant scenario, the measured and unmeasured disturbances were selected to ensure that the recorded variables show a similar variance and pattern compared to that of the industrial examples in Sections 5 and 6. Figure 2 presents the time-based trends of the simulated variables.

[Figure 2 about here]

4.2 Identification of a Dynamic Monitoring Model

From this process, a data set containing a total of 20000 data points was used as a reference set to identify the dynamic models. To assess which of the 28 recorded process variables are nonstationary, we applied the ADF test (Dickey & Fuller, 1981). Table 4 reports the results of the ADF test statistic for which the critical value was determined for a significance of 0.01. This yielded that 27 variables were nonstationary and the other 1 was stationary, *i.e.*, $\mathbf{y}_{\text{ns}}^{(g)} \in \mathbb{R}^{27}$ and $\mathbf{y}_{\text{s}}^{(g)} \in \mathbb{R}$. The stationary variable was the riser temperature, which was adjusted by feedback control. Next, the dimension of nonstationary factors was determined to be $n_{\text{ns}} = 3$ using the Chigira procedure (Chigira, 2008), *i.e.*, $\boldsymbol{\beta}_{\perp}^{(g)\text{T}} \mathbf{y}_{\text{ns}}^{(g)} \in \mathbb{R}^3$. Figure 3 shows the eigenvalues of the sample covariance matrix of the stationary vector and highlights $n_{\text{s}} = 5$. The Chigira procedure yielded the factor loading matrices $\boldsymbol{\beta} \in \mathbb{R}^{28 \times 5}$ and $\boldsymbol{\beta}_{\perp} \in \mathbb{R}^{28 \times 3}$ to be:

$$\boldsymbol{\beta} = \begin{bmatrix} -0.015 & 0.207 & -0.386 & 0.428 & 0.571 \\ -0.110 & 0.161 & -0.053 & 0.070 & -0.152 \\ -0.086 & -0.49 & 0.263 & -0.529 & 0.477 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0.003 & 0.022 & 0.028 & -0.027 & 0.031 \\ 0.026 & 0.046 & 0.006 & -0.023 & -0.115 \\ 0.004 & 0.015 & 0.018 & -0.035 & 0.028 \end{bmatrix}, \quad (18a)$$

$$\boldsymbol{\beta}_{\perp} = \begin{bmatrix} -0.134 & -0.164 & -0.491 \\ 0.080 & 0.282 & -0.207 \\ -0.168 & 0.164 & -0.313 \\ \vdots & \vdots & \vdots \\ 0.263 & -0.035 & -0.030 \\ 0.050 & 0.288 & -0.262 \\ 0.262 & -0.039 & -0.039 \end{bmatrix}. \quad (18b)$$

[Table 4 about here]

[Figure 3 about here]

The next step was to construct the VARI model for the 3 nonstationary factors and the VAR model for the 5 stationary factors. Applying the BIC criteria (Hannan, 1980), the number of lagged term for the VARI and VAR model were estimated to be $p_{ns} = 13$ and $p_s = 6$, respectively. The resultant auto-regressive coefficient matrices for both models are listed in Appendix C.1. The estimated covariance matrices for the model residuals \mathbf{e}_f and \mathbf{e}_g were:

$$\hat{\mathbf{R}}_f = \begin{bmatrix} 0.0003 & -0.0008 & 0.0002 \\ -0.0008 & 0.0059 & -0.0008 \\ 0.0002 & -0.0008 & 0.0002 \end{bmatrix}, \quad (19a)$$

$$\hat{\mathbf{R}}_g = \begin{bmatrix} 0.0002 & -0.0002 & -0.0000 & 0.0003 & -0.0002 \\ -0.0002 & 0.0010 & 0.0001 & -0.0012 & 0.0010 \\ -0.0000 & 0.0001 & 0.0005 & -0.0001 & -0.0006 \\ 0.0003 & -0.0012 & -0.0001 & 0.0015 & -0.0012 \\ -0.0002 & 0.0010 & -0.0006 & -0.0012 & 0.0058 \end{bmatrix}. \quad (19b)$$

4.3 Application of Identified Monitoring Scheme

The estimated dynamic model was applied to a fault scenario. The simulated fault scenario was recorded over a period covering 3000 data points, and the fault was injected 500 data points into the recorded data set. The fault condition is a 1% increase of the coefficient describing the friction in the spent catalyst line, which yields a degeneration in the spent catalyst circulation. The residuals for the data set were produced by the estimated

VARI and VAR models. In addition to that, the compensation scheme introduced in Subsection 3.4 was employed to avoid forecast recovery. Plots (a) and (b) in Figure 4 show the T_{ns}^2 and T_s^2 statistics for the identified nonstationary and stationary factors of the fault scenario, respectively. With respect to the control limits, determined for a significance of 0.01, both plots illustrate that the fault condition was detected 500 data points into the data set.

For comparison, plots (c) and (d) in Figure 4 demonstrates the monitoring performances of a PCA-based monitoring scheme could also detect the fault condition but produced several false alarms during the first 500 samples where the fault had not occurred, compared to the improved common trends framework.

[Figure 4 about here]

More importantly, however, the standard MSPC process monitoring methodology relies on the assumption that the process variables are stationary and hence, it cannot correctly model nonstationary behavior that is embedded within recorded data sets. This is exemplified here by applying a standard PCA model to an additional data set that contains 5000 data points, describing normal process behavior. This set was neither used to identify the VARI and VAR models nor to identify which of the 28 variables were nonstationary. Figure 5 shows the standard nonnegative quadratic statistics based on a PCA monitoring model (Kruger & Xie, 2012) for the additional data set and confirms, as expected, a considerable number of false alarms.

[Figure 5 about here]

4.4 Conclusions

The model, based on the improved common trends framework, introduced in Section 3, has shown a substantially better performance for monitoring the FCCU compared to a conventional MSPC-based approach. This is because data structure in Eq. (15) and its corresponding Kasa decomposition is designed specifically to handle nonstationary and serially correlated process variables. Conversely, conventional MSPC methods can handle serial correlation but assume stationary variables, which has led to the considerable number of false alarms according to Figure 5. Concerning the improvements reported in Subsection 3.5, this application has confirmed:

- the need to invoke the Chigira procedure, given that there are no critical values for $n = 27$ nonstationary variables – for practical applications, the Chigira procedure is to be generally preferred, as it is not compromised by the number of nonstationary variables; and

- the need to utilize the generalized data structure in Eq. (15), as the recorded variable set contained stationary and nonstationary variables.

In addition to that, this application study has confirmed that differencing the nonstationary variables once produced stationary sequences, which has highlighted that the assumption $d = 1$ is justifiable in practice. For all application studies, Table 8 summarizes the results and confirms that, as far as the detection of the fault is concerned, the common trends framework is as successful as a conventional PCA approach. However, the false alarm rate (FAR) of the PCA-based monitoring approach is unacceptable. In addition to that, the missing alarm rate (MAR) is defined here as the number of instances correctly detected as abnormal over the total number of data points that describe the fault condition.

5 Application to the Polymerization Process

This section applies the improved common trends framework to the industrial polymerization process, described in Subsection 2.2. Details on how the monitoring model was established are given in Subsection 5.1. This is followed by applying the monitoring model to detect an abnormal event in Subsection 5.2. Finally, subsection 5.3 concludes this application study. Figure 6 presents time-based plots for the 7 recorded variables to highlight that some variables are stationary, whilst others show nonstationary trends (y_1 : gas flow, y_2 : positional gas heater temperature and y_4 : head metal temperature).

[Figure 6 about here]

5.1 Identification of a Dynamic Monitoring Model

For this process, a reference data set containing a total of 10000 data points was used to identify the dynamic models. Table 5 reports the results of applying the ADF test (Dickey & Fuller, 1981), which yielded $\mathbf{y}_{\text{ns}}^{(g)} \in \mathbb{R}^3$ and $\mathbf{y}_s^{(g)} \in \mathbb{R}^4$. Moreover, applying the Chigira and Johansen procedures both revealed that $\boldsymbol{\beta}_{\perp}^{(g)\text{T}} \mathbf{y}_{\text{ns}}^{(g)} \in \mathbb{R}^2$ and $\boldsymbol{\beta}^{(g)\text{T}} \mathbf{y}_{\text{ns}}^{(g)} \in \mathbb{R}$. Whilst $n_{\text{ns}} = 2$, the Chigira procedure also yielded $n_s = 5$ and produced the following factor loading matrices $\boldsymbol{\beta} \in \mathbb{R}^{7 \times 5}$ and $\boldsymbol{\beta}_{\perp} \in \mathbb{R}^{7 \times 2}$:

$$\boldsymbol{\beta} = \begin{bmatrix} -0.007 & 0.017 & 0.386 & 0.423 & -0.081 \\ -0.010 & 0.023 & 0.544 & 0.595 & -0.114 \\ -0.368 & -0.641 & -0.478 & 0.470 & 0.072 \\ 0.000 & -0.001 & -0.018 & -0.019 & 0.004 \\ -0.489 & -0.483 & 0.532 & -0.490 & -0.074 \\ -0.559 & 0.426 & 0.039 & 0.074 & 0.707 \\ -0.560 & 0.417 & -0.204 & 0.029 & -0.686 \end{bmatrix}, \quad \boldsymbol{\beta}_{\perp} = \begin{bmatrix} -0.770 & 0.270 \\ 0.557 & -0.161 \\ 0 & 0 \\ 0.313 & 0.949 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (20)$$

[Table 5 about here]

The next step was to construct the VARI model for the 2 nonstationary and the VAR model for the 5 stationary factors. Applying the BIC index (Hannan, 1980), the number of lagged term for the VARI and VAR model were estimated to be $p_{ns} = 12$ and $p_s = 9$, respectively. The resultant auto-regressive coefficient matrices for both models are listed in Appendix C.2. The estimated covariance matrices for the model residuals \mathbf{e}_f and \mathbf{e}_g were:

$$\hat{\mathbf{R}}_f = \begin{bmatrix} 0.122 & -0.001 \\ -0.001 & 0.183 \end{bmatrix} \quad \hat{\mathbf{R}}_g = \begin{bmatrix} 0.850 & 0.031 & -0.164 & 0.021 & 0.012 \\ 0.031 & 0.803 & 0.222 & -0.028 & 0.051 \\ -0.164 & 0.222 & 0.561 & 0.091 & -0.055 \\ 0.021 & -0.028 & 0.091 & 0.593 & 0.183 \\ 0.012 & 0.051 & -0.055 & 0.183 & 0.879 \end{bmatrix}. \quad (21)$$

The application of the Anderson-Darling or AD test (Anderson & Darling, 1954) confirmed, for a significance of $\alpha = 0.01$, that the calculated residual sequences were drawn from multivariate normal distributions.

5.2 Application of Identified Monitoring Model

The recorded data also contains a segment containing 5000 data points that describe a fault condition that affected variables y_3 and y_4 . The root cause for this fault condition could not be determined but it was confirmed that the anomalous event affected the quality of the threads. Figure 7 outlines that the fault condition emerged 600 to 650 data points into the data set. As discussed in Section 3, the improved common trends framework scheme relied on the estimated VARI and VAR models, and the compensation scheme to avoid the problem of forecast recovery. Plots (a) and (b) in Figure 8 show the T_{ns}^2 and T_s^2 statistics describing the variation of the identified nonstationary and stationary factors, respectively. With respect to the control limits that were determined for a significance of 0.01, both plots show that the fault condition was detected around 600 samples into the data set.

[Figure 7 about here]

For comparison, a standard MSPC approach, based on PCA, was also applied. Plots (c) and (d) in Figure 8 display the corresponding Hotelling's T^2 and SPE statistics, which produced a considerable number of false

alarms prior to the occurrence of the fault condition. This is not surprising, given that the assumptions imposed on conventional MSPC do not allow modeling nonstationary variables, as highlighted in the previous section. In addition to that, the PCA-based monitoring model provided a sporadic detection of this abnormal. To examine this further, we recomputed the monitoring statistics based on the improved common trends model but without applying the compensation scheme. Plots (a) and (b) in Figure 9 show the recomputed T_{ns}^2 and T_s^2 statistics. Comparing plots (a) and (b) in Figures 8 and 9 highlights that the compensation scheme yields a fault signature that is consistent with the effect it had on the product quality.

[Figure 8 about here]

[Figure 9 about here]

5.3 Conclusions

This industrial application study has shown that the polymerization process requires the use of the improved common trends framework to monitor the seven recorded variables. Similar to the simulated study in the previous section, although conventional MSPC approach can deal with serial correlation that is embedded within the recorded data it is not capable of modeling the three nonstationary variables y_1 , y_2 and y_4 . The diagnosis of the detected fault condition has confirmed that extracting the fault signature required the use of the compensation scheme, as the original residuals do only show sporadic violations that did not reflect the undesired impact of this event upon the product quality. The application of the common trends framework to this process has required the following improvements, detailed in Subsection 3.5:

- given that the variable set contained stationary and nonstationary variables, the generalized data structure in Eq. (15) had to be utilized; and
- although critical values for the Johansen procedure are available for $n = 3$ nonstationary variables, the Chigira procedure produced an identical estimate for the number of nonstationary factors; and the Chigira procedure is conceptually simpler than the Johansen procedure and closely related to the MSPC framework.

By applying the Anderson-Darling test, not shown here, the residual sequences of the VARI and VAR models were drawn from multivariate normal distributions. Finally, differencing the three nonstationary variables has produced stationary sequences and hence, this industrial application study has also confirmed that the assumption $d = 1$ has been met. Table 8 summarizes the above results and also confirm, as observed in Figures 8 and 9, that the standard PCA approach produced false alarms, particularly the T^2 statistic, and had

very high levels of missed alarms, which are around 40% for the T^2 statistic and in excess of 98% for the SPE statistic. Table 8 also outlines the importance of the compensation procedure.

6 Application to the Distillation Unit

This section applied the introduced monitoring scheme to an industrial instance of a distillation unit that is described in Subsection 2.3. Details of identifying the monitoring model are given in Subsection 6.1. This is followed by an application to detect an abnormal process event in Subsection 6.2. Finally, Subsection 6.3 concludes this industrial application. Figure 10 presents time-based plots for the 12 recorded variables to highlight that some variables are stationary, whilst others show nonstationary trends (y_3 : tray 2 temperature, y_6 : reflux flow, y_8 : butane product flow and y_{12} : reboiler outlet temperature).

[Figure 10 about here]

6.1 Identification of a Dynamic Monitoring Model

To model the distillation process, a reference set containing a total of 10000 data points was used to identify the monitoring model. Applying the ADF test (Dickey & Fuller, 1981) to each of the 12 recorded variables, for a significance of 0.01, yielded that 4 of them are nonstationary, *i.e.*, y_3 , y_6 , y_8 and y_{12} , whilst the remaining 8 are stationary. Table 6 lists the results of the ADF tests. A further analysis of the 4 nonstationary variables using the Johansen and Chigira procedures revealed $n_{ns} = 2$ nonstationary and $n_s = 10$ stationary factors, and extracted the following factor loading matrices $\beta \in \mathbb{R}^{12 \times 10}$ and $\beta_{\perp} \in \mathbb{R}^{12 \times 2}$:

$$\beta = \begin{bmatrix} -0.157 & -0.661 & 0.128 & \cdots & -0.098 & 0.027 \\ 0.051 & 0.005 & 0.635 & \cdots & -0.039 & -0.012 \\ 0.002 & 0.003 & 0.020 & \cdots & -0.263 & 0.011 \\ 0.055 & 0.185 & 0.302 & \cdots & -0.018 & -0.000 \\ 0.553 & -0.026 & 0.170 & \cdots & 0.237 & 0.082 \\ -0.004 & 0.003 & 0.012 & \cdots & -0.352 & 0.726 \\ 0.059 & -0.719 & 0.002 & \cdots & 0.119 & -0.022 \\ 0.011 & 0.006 & 0.043 & \cdots & -0.375 & -0.677 \\ 0.021 & 0.015 & 0.641 & \cdots & -0.024 & -0.004 \\ 0.553 & -0.103 & -0.208 & \cdots & -0.384 & -0.055 \\ 0.595 & -0.001 & -0.037 & \cdots & 0.116 & 0.002 \\ -0.006 & -0.008 & -0.049 & \cdots & 0.650 & -0.048 \end{bmatrix}, \quad \beta_{\perp} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -0.447 & 0.842 \\ 0 & 0 \\ 0 & 0 \\ 0.551 & 0.148 \\ 0 & 0 \\ 0.553 & 0.138 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0.437 & 0.500 \end{bmatrix}. \quad (22)$$

[Table 6 about here]

The next step was identifying the VARI model for the 2 nonstationary factors and the VAR model for the 10 stationary factors. By applying the BIC index (Hannan, 1980), the number of lagged terms for the VARI and VAR model were estimated to be $p_{ns} = 4$ and $p_s = 7$, respectively. The resultant auto-regressive coefficient matrices for both models are listed in Appendix C.3. The sample covariance matrices for the model residuals \mathbf{e}_f and \mathbf{e}_g were:

$$\hat{\mathbf{R}}_f = \begin{bmatrix} 0.003 & 0.001 \\ 0.001 & 0.008 \end{bmatrix}, \quad \hat{\mathbf{R}}_g = \begin{bmatrix} 0.020 & 0.004 & 0.029 & \cdots & 0.028 & 0.016 \\ 0.004 & 0.050 & 0.066 & \cdots & -0.042 & 0.010 \\ 0.029 & 0.066 & 0.430 & \cdots & -0.023 & 0.016 \\ -0.000 & -0.018 & -0.127 & \cdots & 0.005 & 0.017 \\ 0.002 & -0.001 & -0.003 & \cdots & 0.017 & -0.038 \\ 0.015 & 0.022 & -0.009 & \cdots & 0.017 & 0.012 \\ -0.009 & 0.003 & 0.045 & \cdots & -0.016 & -0.068 \\ 0.001 & 0.015 & -0.026 & \cdots & -0.043 & 0.071 \\ 0.028 & -0.042 & -0.023 & \cdots & 0.702 & 0.041 \\ 0.016 & 0.010 & 0.016 & \cdots & 0.041 & 0.499 \end{bmatrix}. \quad (23)$$

6.2 Application of Identified Monitoring Model

The recorded data set also contained a series of severe drops in fresh feeds. This section of the data set contained 5000 data points and described two drops in feed that arose after around 1100 and 3000 points into this data set. The plant operators responded to the presence of the first drop, but did not notice the second feed drop. Without operator intervention, a prolonged drop in fresh feed upsets the diffusion conditions within the column, which is predominantly noticeable by an increase in the measured temperatures.

The residuals, computed by the estimated VARI and VAR models, and the application of the compensation scheme, introduced in Subsection 3.4, to avoiding the problem of forecast recovery, were then used to calculate the T_{ns}^2 and T_s^2 statistics for each data point. The results of applying the Anderson-Darling or AD test (Anderson & Darling, 1954) to the calculated statistics to the residuals of the VARI and VAR models, listed in Table 7 for a significance of 0.01, showed that some residual sequences were not drawn from normal distributions. This required estimating the probability density function of the T_{ns}^2 and T_s^2 statistics in order to numerically compute their control limits, as described in Subsection 3.5.3.

[Table 7 about here]

Plots (a) and (b) in Figure 11 show time based plots of the T_{ns}^2 and T_s^2 statistics, respectively. Plots (a) and (b) were briefly sensitive to the presence of the first drop at around 1100th sample, although this event lasted only for a short period. The stationary T_s^2 statistic in plot (b) had a substantially response to the second

drop in feed around 3000 data points into the data set. In contrast, the T_{ns}^2 statistic in plot (a) became significant to the second feed drop around 3300 points into the data set. The fresh feed, *i.e.*, variable y_{11} , was characterized as a stationary variable. Any significant and abnormal variation of y_{11} relative to the remaining stationary variables or the stationary components of the nonstationary variables must manifest itself in the residuals of the VAR model and, hence, the T_s^2 statistic.

[Figure 11 about here]

Thermodynamically, a prolonged drop in fresh feed upsets the diffusion conditions within the column, which, in turn, impacts the column temperatures first and subsequently the flow rates. Consequently, the effect of the feed drop to the relationship of the nonstationary butane product flow relative to the other nonstationary variables was felt with some delay by the T_{ns}^2 statistic, whilst the drop in feed affected the temperatures almost instantly which resulted in a significant response by the T_s^2 statistic.

For comparison, a conventional MSPC approach, based on PCA, was also applied to model the data for this distillation unit. Plots (c) and (d) in Figure 11 show the monitoring performance based on a PCA-based monitoring statistics. As witnessed in the previous two application studies, plots (c) and (d) both show a significant number of false alarms prior to the second and more severe and prolonged drop in fresh feed. This is not surprising, given that a standard MSPC approach assumes that the recorded process variables are stationary.

We finally examined the impact of violating the assumption that the calculated residual samples were drawn from multivariate normal distributions. Plots (a) and (b) in Figure 12 show the T_{ns}^2 and T_s^2 monitoring statistics, constructed from a third data set, containing 3000 data points, and describing normal process operation when benchmarked against the control limits computed from Eqs. (8) and (11), respectively. The T_{ns}^2 statistic, constructed by two residual variables of which only one has a normal distribution, did not show an unexpectedly large number of false alarms. In fact, the relative number of violations did not exceed the significance of 0.01. However, a different picture emerged when benchmarking the T_s^2 statistic against its control limit in plot (b). In this case, the relative number of violations well exceeded the significance of 0.01 and implied that the control limit was calculated to be too small. Hence, the considerably larger control limit, computed with respect to the estimated PDF, was required, which highlights the importance of the improvement detailed in Subsection 3.5.3.

[Figure 12 about here]

6.3 Conclusions

This application study has also shown that the improved common trends framework has been required for monitoring the distillation process. With 4 out of the recorded 12 process variables being nonstationary, the application of PCA has shown that the MSPC methodology is not capable to model such behavior, although the two drops in fresh feed were noticeable in both statistics. To make the common trends framework applicable to this process has required incorporating the following improvements, detailed in Subsection 3.5:

- the presence of stationary and nonstationary variables required the use of the generalized data structure in Eq. (15);
- although critical values for the Johansen procedure are available for $n_{ns} = 4$ nonstationary variables, the Chigira procedure produced an identical estimate for the number of nonstationary factors; and the Chigira procedure is conceptually simpler than the Johansen procedure and closely related to the MSPC framework; and
- a number of residuals sequences were not drawn from normal distributions, which required the use of estimated probability density functions in order to determine the control limits.

The effect of incorrectly assuming that the control limits can be computed from Eqs. (8) and (11) has also been demonstrated and has yielded that, particularly for the T_s^2 statistic, it has the potential to produce false alarms. Finally, differencing the 4 nonstationary variables has produced stationary sequences, which confirm that this industrial application study could also rely on the assumption $d = 1$. These results are also summarized in Table 8. In addition to that, Table 8 also summarizes the percentage of false alarms by the traditional PCA-based monitoring approach, which exceeds 40% for the T^2 statistic and 90% for the SPE statistic. Table 8 also confirms that relying on the assumption that the residuals of the VAR and VARI models were drawn from a multivariate normal distribution can produce a false alarm rate of over 5% although the significance was selected to be 1% only.

[Table 8 about here]

7 Concluding Summary

This article has introduced practically important improvements to a recently proposed common trends framework for monitoring industrial production systems that yield process variables that are serially correlated and nonstationary. In industrial practice, this is a typical scenario and result from controller feedback, changes in process feeds, unmeasured disturbances etc. The improvements that have been discussed here (i) include the development of a generalized data structure to model variable sets that contain stationary as well as nonstationary variables, (ii) highlight that the standard Johansen method for identifying cointegration models is not

suitable for monitoring applications in a general context and (iii) outline how to handle model residual sequences that are not drawn from multivariate normal distributions. The paper has argued for the need of these improvements on the basis of three application studies that include the simulation of a fluid catalytic cracking unit and the analysis of two industrial data sets that involve a spinning process and a distillation unit.

To be able to construct a common trends model, the paper has advocated the use of the Chigira procedure instead of the Johansen method. The Chigira procedure is closely related to the MSPC framework and not compromised by the problems of the Johansen method, *i.e.*, the lack of empirical confidence limits for larger variable sets and the need of pre-identified auto-regressive models for each nonstationary variable. To the best of our knowledge, such a detailed practical comparison for different methods to obtain cointegration models has not been presented in the literature. The paper also has compared the improved common trends framework with a conventional MSPC approach. This comparison has shown that the standard MSPC approach is unable to adequately model the nonstationary data, whilst the improved common trends framework has not resulted in the production of false alarms. In addition to that, the improved common trends framework has been able to diagnose two simulated fault conditions for the FCCU and the recorded fault conditions for the spinning process and the distillation unit.

Acknowledgement

The authors gratefully acknowledge support from a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

Appendix A: Vector Auto-Regressive Integrated Models

The stochastic vector \mathbf{y} is assumed to be integrated of order $d = 1$ and does not have a cointegration relationship. That is, its difference $\nabla \mathbf{y}(k) \equiv \mathbf{y}(k) - \mathbf{y}(k - 1)$ is stationary and can be represented in the form of a VAR model:

$$\nabla \mathbf{y}(k) = \sum_{j=1}^p \Theta_j \nabla \mathbf{y}(k - j) + \mathbf{e}(k), \quad (\text{A1})$$

where p is the number of lagged terms that can be estimated using the Bayesian information criteria, or BIC (Hannan, 1980), $\Theta_j, j = 1, \dots, p_{\text{ns}}$, are coefficient matrices, and \mathbf{e} denotes a random error vector that has a multivariate normal distributions with a mean of zero and the covariance matrix \mathbf{R} . Eq. (A1) can be reformulated in the form of a VARI model for random vector \mathbf{y}

$$\mathbf{y}(k) = \sum_{j=1}^{p+1} \mathbf{\Psi}_j \mathbf{y}(k-j) + \mathbf{e}(k), \quad (\text{A2})$$

where the coefficient matrices $\mathbf{\Psi}_j$, $j = 1, \dots, p, p+1$ are given by

$$\mathbf{\Psi}_j = \begin{cases} \mathbf{I} + \mathbf{\Theta}_1 & \text{if } j = 1 \\ \mathbf{\Theta}_j - \mathbf{\Theta}_{j-1} & \text{if } j = 2, \dots, p. \\ -\mathbf{\Theta}_{p+1} & \text{if } j = p+1 \end{cases} \quad (\text{A3})$$

Appendix B: Augmented Dickey-Fuller Test

A recorded serially correlated sequence is said to be stationary if its mean and auto-correlation function are time-invariant; otherwise the sequence follows a nonstationary process. The auto-correlation function is a useful tool for analyzing such recorded sequences where a set of data points, recorded within a small time window, are statistically dependent upon each another. Each data point in a recorded set is modeled as a linear combination of the previous records to which an element of excitation noise from a random innovation process is superimposed. Generally, a p -order auto-regressive model is of the following form:

$$y(k) = a_1 y(k-1) + a_2 y(k-2) + \dots + a_p y(k-p) + e(k), \quad (\text{B1})$$

where $y(k)$ denotes to the k th data point of the recorded process variable y , a_i ($i = 1, \dots, p$) are model parameters, and e represents a stationary serially uncorrelated variable with zero mean and variance σ^2 . The characteristic equation of Eq. (B1) is $1 - a_1 z - a_2 z^2 - \dots - a_p z^p = 0$, and if the absolute value of all its roots is less than 1, the sequence $\{y(k)\}$ is stationary, otherwise, $\{y(k)\}$ is nonstationary. Thus, the test for nonstationarity reduces to a test for the presence of unit roots, namely, transferring the nonstationary test to verify the presence of at least one unit root.

Detecting the presence of the unit root for each recorded process variable is the starting point of nonstationary process modeling. Many approaches for testing the presence of unit roots in serial sequences have been proposed (Dickey & Fuller, 1981; Phillips & Perron, 1988; Fan & Gencay, 2010). The Augmented Dickey-Fuller or ADF test (Dickey & Fuller, 1981), has obtained the widest application in practice over the past decades due to its simplicity and reliability, and a brief outline of the ADF test procedure is given below.

The ADF test constructs a parametric correction for higher-order correlation by assuming that the recorded serial sequence $\{y(k)\}$ follows an auto-regressive process with an order of p and adding p lagged differenced terms of the data points $y(k-i)$, ($i = 1, 2, \dots, p$) to the right side of the testing equation as:

$$\Delta y(k) = \rho y(k-1) + \sum_{i=1}^p a_i \Delta y(k-i) + e(k), \quad (\text{B2})$$

after estimating the parameters of the model in Eq. (B2), the null and alternative hypothesis include $H_0: \rho = 0$; $H_1: \rho < 0$. If the null hypothesis H_0 , where the critical values of the t statistic were tabulated by Dickey and Fuller, fails to be rejected, it must be concluded that there is a unit root. More details about the testing procedure can be found in Pfaff (2008).

Appendix C: Identified Time Series Models for Industrial Processes

C.1 Identified Parameter Matrices of VARI and VAR Models for FCCU Application

The identified dynamic VARI model for the nonstationary factors of the FCCU has the following parameter matrices:

$$\mathbf{f}(k) = \sum_{i=1}^{14} \Psi_i \mathbf{f}(k-i) + \mathbf{e}_f(k), \quad (\text{C1})$$

where:

$$\Psi_1 = \begin{bmatrix} 1.658 & 0.001 & 0.080 \\ 1.478 & 1.491 & 0.357 \\ -0.439 & 0.009 & 1.556 \end{bmatrix}, \quad (\text{C2a})$$

$$\Psi_2 = \begin{bmatrix} -0.946 & 0.010 & -0.075 \\ -1.823 & -0.667 & -0.556 \\ 0.566 & -0.031 & -0.661 \end{bmatrix}, \quad (\text{C2b})$$

$$\Psi_3 = \begin{bmatrix} 0.462 & -0.013 & -0.001 \\ 0.539 & 0.333 & 0.219 \\ -0.179 & 0.021 & 0.338 \end{bmatrix}, \quad (\text{C2c})$$

$$\Psi_4 = \begin{bmatrix} -0.250 & 0.005 & -0.025 \\ -0.312 & -0.191 & -0.092 \\ 0.102 & -0.000 & -0.207 \end{bmatrix}, \quad (\text{C2d})$$

$$\Psi_5 = \begin{bmatrix} 0.129 & -0.000 & 0.021 \\ 0.143 & 0.071 & -0.035 \\ -0.045 & 0.011 & 0.092 \end{bmatrix}, \quad (\text{C2e})$$

$$\Psi_6 = \begin{bmatrix} -0.062 & 0.002 & -0.015 \\ -0.115 & -0.055 & 0.033 \\ 0.028 & 0.001 & -0.082 \end{bmatrix}, \quad (\text{C2f})$$

$$\Psi_7 = \begin{bmatrix} 0.027 & -0.004 & -0.011 \\ 0.153 & 0.033 & 0.081 \\ -0.036 & -0.003 & -0.008 \end{bmatrix}, \quad (\text{C2g})$$

$$\Psi_8 = \begin{bmatrix} -0.029 & -0.001 & 0.012 \\ -0.111 & -0.016 & -0.033 \\ 0.016 & 0.002 & -0.013 \end{bmatrix}, \quad (\text{C2h})$$

$$\Psi_9 = \begin{bmatrix} 0.011 & -0.001 & 0.023 \\ 0.072 & -0.002 & -0.097 \\ -0.011 & 0.000 & 0.013 \end{bmatrix}, \quad (\text{C2i})$$

$$\Psi_{10} = \begin{bmatrix} 0.012 & -0.003 & -0.038 \\ -0.074 & 0.023 & 0.196 \\ 0.020 & -0.006 & -0.050 \end{bmatrix}, \quad (\text{C2j})$$

$$\Psi_{11} = \begin{bmatrix} 0.014 & 0.006 & 0.005 \\ -0.004 & -0.028 & -0.055 \\ 0.002 & 0.007 & 0.017 \end{bmatrix}, \quad (\text{C2k})$$

$$\Psi_{12} = \begin{bmatrix} -0.056 & -0.004 & 0.026 \\ 0.112 & 0.007 & -0.075 \\ -0.026 & -0.009 & -0.022 \end{bmatrix}, \quad (\text{C2l})$$

$$\Psi_{13} = \begin{bmatrix} 0.047 & -0.003 & -0.024 \\ -0.095 & 0.011 & 0.130 \\ -0.001 & 0.000 & 0.005 \end{bmatrix}, \quad (\text{C2m})$$

$$\Psi_{14} = \begin{bmatrix} -0.017 & 0.005 & 0.021 \\ 0.035 & -0.011 & -0.071 \\ 0.003 & -0.002 & 0.021 \end{bmatrix}. \quad (\text{C2n})$$

The identified dynamic VAR model for the stationary factors of the FCCU has the following parameter matrices:

$$\mathbf{g}(k) = \sum_{i=1}^6 \Phi_i \mathbf{g}(k-i) + \mathbf{e}_g(k), \quad (\text{C3})$$

where:

$$\Phi_1 = \begin{bmatrix} 1.314 & 0.160 & -0.030 & 0.130 & 0.029 \\ 0.153 & 1.143 & 0.062 & -0.167 & -0.106 \\ -0.002 & -0.030 & 1.455 & -0.026 & -0.013 \\ -0.225 & 0.865 & -0.121 & 2.115 & 0.044 \\ 0.953 & -3.910 & 0.397 & -2.861 & 1.476 \end{bmatrix}, \quad (\text{C4a})$$

$$\Phi_2 = \begin{bmatrix} -0.471 & -0.218 & 0.033 & -0.211 & -0.021 \\ -0.032 & -0.688 & -0.017 & -0.068 & 0.097 \\ 0.066 & -0.144 & -0.651 & -0.106 & 0.023 \\ 0.027 & -0.458 & 0.062 & -1.078 & 0.027 \\ -0.535 & 3.255 & -0.231 & 2.532 & -1.047 \end{bmatrix}, \quad (\text{C4b})$$

$$\Phi_3 = \begin{bmatrix} 0.196 & 0.154 & -0.016 & 0.175 & -0.003 \\ -0.014 & 0.472 & -0.022 & 0.119 & -0.045 \\ -0.086 & 0.207 & 0.304 & 0.174 & -0.031 \\ 0.072 & -0.130 & 0.024 & 0.280 & -0.009 \\ -0.003 & -0.323 & -0.109 & -0.437 & 0.548 \end{bmatrix}, \quad (\text{C4c})$$

$$\Phi_4 = \begin{bmatrix} -0.075 & -0.075 & 0.012 & -0.095 & 0.001 \\ -0.028 & -0.219 & 0.014 & -0.062 & 0.036 \\ 0.038 & -0.057 & -0.158 & -0.059 & 0.025 \\ 0.019 & 0.023 & -0.011 & -0.149 & -0.028 \\ -0.129 & 0.122 & 0.132 & 0.120 & -0.146 \end{bmatrix}, \quad (\text{C4d})$$

$$\Phi_5 = \begin{bmatrix} 0.051 & -0.029 & -0.003 & -0.010 & -0.002 \\ -0.071 & 0.374 & -0.031 & 0.268 & -0.005 \\ -0.020 & 0.025 & 0.082 & 0.018 & -0.007 \\ 0.101 & -0.359 & 0.034 & -0.235 & -0.003 \\ -0.238 & 0.898 & -0.162 & 0.784 & 0.097 \end{bmatrix}, \quad (\text{C4e})$$

$$\Phi_6 = \begin{bmatrix} -0.016 & 0.007 & 0.004 & 0.010 & -0.003 \\ -0.007 & -0.082 & -0.005 & -0.087 & 0.015 \\ 0.005 & -0.002 & -0.032 & -0.001 & 0.003 \\ 0.006 & 0.058 & 0.012 & 0.062 & -0.021 \\ -0.049 & -0.042 & -0.026 & -0.130 & 0.042 \end{bmatrix}. \quad (\text{C4f})$$

C.2 Identified Parameter Matrices of VARI and VAR Models for Polymerization Process

The identified dynamic VARI model for the nonstationary factors of the polymerization process has the following parameter matrices:

$$\mathbf{f}(k) = \sum_{i=1}^{13} \Psi_i \mathbf{f}(k-i) + \mathbf{e}_f(k), \quad (\text{C5})$$

where:

$$\Psi_1 = \begin{bmatrix} 0.509 & -0.064 \\ -0.086 & 0.212 \end{bmatrix}, \quad (\text{C6a})$$

$$\Psi_2 = \begin{bmatrix} 0.030 & 0.011 \\ 0.021 & 0.174 \end{bmatrix}, \quad (\text{C6b})$$

$$\Psi_3 = \begin{bmatrix} 0.243 & -0.024 \\ 0.003 & 0.134 \end{bmatrix}, \quad (\text{C6c})$$

$$\Psi_4 = \begin{bmatrix} -0.498 & 0.078 \\ 0.099 & 0.101 \end{bmatrix}, \quad (\text{C6d})$$

$$\Psi_5 = \begin{bmatrix} 0.220 & -0.018 \\ -0.004 & 0.100 \end{bmatrix}, \quad (\text{C6e})$$

$$\Psi_6 = \begin{bmatrix} 0.175 & 0.006 \\ -0.023 & 0.062 \end{bmatrix}, \quad (\text{C6f})$$

$$\Psi_7 = \begin{bmatrix} 0.012 & 0.019 \\ 0.013 & 0.068 \end{bmatrix}, \quad (\text{C6g})$$

$$\Psi_8 = \begin{bmatrix} 0.028 & -0.002 \\ -0.002 & 0.038 \end{bmatrix}, \quad (\text{C6h})$$

$$\Psi_9 = \begin{bmatrix} 0.082 & -0.004 \\ -0.013 & 0.029 \end{bmatrix}, \quad (\text{C6i})$$

$$\Psi_{10} = \begin{bmatrix} 0.078 & -0.001 \\ 0.007 & 0.039 \end{bmatrix}, \quad (\text{C6j})$$

$$\Psi_{11} = \begin{bmatrix} -0.003 & 0.004 \\ -0.000 & 0.014 \end{bmatrix}, \quad (\text{C6k})$$

$$\Psi_{12} = \begin{bmatrix} 0.004 & -0.004 \\ -0.002 & -0.016 \end{bmatrix}, \quad (\text{C6l})$$

$$\Psi_{13} = \begin{bmatrix} 0.104 & 0.003 \\ -0.017 & 0.019 \end{bmatrix}. \quad (\text{C6m})$$

The identified dynamic VAR model for the stationary factors of the polymerization process has the following parameter matrices:

$$\mathbf{g}(k) = \sum_{i=1}^9 \Phi_i \mathbf{g}(k-i) + \mathbf{e}_g(k), \quad (\text{C7})$$

where:

$$\Phi_1 = \begin{bmatrix} 0.074 & 0.011 & -0.027 & -0.001 & -0.006 \\ -0.014 & 0.104 & -0.111 & 0.136 & -0.069 \\ 0.047 & -0.031 & 0.173 & -0.233 & 0.130 \\ -0.020 & -0.176 & 0.118 & 0.498 & -0.202 \\ 0.012 & 0.067 & -0.041 & -0.230 & 0.152 \end{bmatrix}, \quad (\text{C8a})$$

$$\Phi_2 = \begin{bmatrix} 0.095 & 0.018 & 0.020 & -0.029 & -0.003 \\ 0.018 & 0.051 & -0.005 & 0.088 & -0.057 \\ 0.038 & -0.005 & 0.095 & -0.050 & 0.040 \\ 0.021 & -0.057 & 0.075 & -0.042 & 0.054 \\ -0.007 & 0.018 & -0.040 & 0.004 & 0.048 \end{bmatrix}, \quad (\text{C8b})$$

$$\Phi_3 = \begin{bmatrix} 0.079 & -0.024 & 0.056 & -0.009 & -0.007 \\ 0.019 & 0.047 & 0.007 & -0.023 & 0.006 \\ 0.026 & -0.033 & 0.105 & 0.009 & 0.001 \\ -0.022 & 0.019 & -0.070 & 0.182 & -0.066 \\ 0.013 & -0.057 & 0.031 & -0.032 & 0.057 \end{bmatrix}, \quad (\text{C8c})$$

$$\Phi_4 = \begin{bmatrix} 0.078 & -0.001 & 0.025 & 0.010 & -0.013 \\ 0.007 & 0.072 & -0.011 & 0.008 & -0.003 \\ 0.011 & -0.034 & 0.078 & 0.057 & -0.019 \\ 0.027 & 0.012 & 0.037 & -0.353 & 0.190 \\ -0.003 & -0.038 & -0.020 & 0.193 & -0.067 \end{bmatrix}, \quad (\text{C8d})$$

$$\Phi_5 = \begin{bmatrix} 0.043 & 0.001 & 0.030 & 0.001 & -0.021 \\ 0.002 & 0.055 & 0.006 & -0.042 & 0.015 \\ 0.016 & -0.035 & 0.090 & 0.025 & -0.015 \\ -0.030 & 0.013 & -0.040 & 0.190 & -0.062 \\ -0.028 & 0.007 & 0.005 & -0.057 & 0.082 \end{bmatrix}, \quad (\text{C8e})$$

$$\Phi_6 = \begin{bmatrix} 0.052 & 0.002 & -0.009 & 0.000 & 0.004 \\ 0.013 & 0.046 & -0.032 & -0.014 & -0.018 \\ 0.014 & -0.023 & 0.068 & 0.003 & -0.001 \\ -0.010 & 0.026 & -0.029 & 0.075 & -0.017 \\ -0.008 & -0.043 & 0.017 & -0.028 & 0.060 \end{bmatrix}, \quad (\text{C8f})$$

$$\Phi_7 = \begin{bmatrix} 0.066 & 0.005 & 0.005 & -0.007 & -0.008 \\ 0.008 & 0.069 & -0.048 & 0.006 & -0.022 \\ 0.008 & -0.017 & 0.057 & 0.045 & -0.018 \\ -0.009 & 0.022 & -0.002 & -0.033 & 0.019 \\ -0.018 & -0.036 & 0.005 & 0.042 & -0.021 \end{bmatrix}, \quad (\text{C8g})$$

$$\Phi_8 = \begin{bmatrix} 0.051 & 0.002 & 0.040 & -0.028 & -0.002 \\ 0.007 & 0.066 & -0.029 & -0.046 & 0.008 \\ 0.021 & -0.017 & 0.046 & 0.002 & -0.000 \\ -0.007 & 0.007 & -0.011 & 0.034 & 0.003 \\ -0.023 & 0.013 & -0.006 & 0.028 & 0.006 \end{bmatrix}, \quad (\text{C8h})$$

$$\Phi_9 = \begin{bmatrix} 0.048 & 0.007 & 0.030 & 0.008 & 0.009 \\ 0.029 & 0.065 & -0.009 & -0.060 & 0.002 \\ 0.016 & -0.020 & 0.038 & 0.035 & -0.027 \\ -0.016 & 0.027 & -0.050 & 0.036 & -0.014 \\ 0.016 & -0.020 & 0.006 & -0.044 & 0.030 \end{bmatrix}. \quad (\text{C8i})$$

C.3 Identified Parameter Matrices of VARI and VAR Models for Distillation Unit

The identified dynamic VARI model for the nonstationary factors of the distillation unit has the following parameter matrices:

$$\mathbf{f}(k) = \sum_{i=1}^5 \Psi_i \mathbf{f}(k-i) + \mathbf{e}_f(k), \quad (\text{C9})$$

where:

$$\Psi_1 = \begin{bmatrix} 0.491 & 0.066 \\ 0.175 & 0.597 \end{bmatrix}, \quad (\text{C10a})$$

$$\Psi_2 = \begin{bmatrix} 0.252 & -0.031 \\ -0.081 & 0.162 \end{bmatrix}, \quad (\text{C10b})$$

$$\Psi_3 = \begin{bmatrix} 0.124 & -0.028 \\ -0.086 & 0.118 \end{bmatrix}, \quad (\text{C10c})$$

$$\Psi_4 = \begin{bmatrix} 0.086 & -0.005 \\ 0.005 & 0.060 \end{bmatrix}, \quad (\text{C10d})$$

$$\Psi_5 = \begin{bmatrix} 0.047 & -0.001 \\ -0.013 & 0.062 \end{bmatrix}. \quad (C10e)$$

The identified dynamic VAR model for the stationary factors of the distillation unit has the following parameter matrices:

$$\mathbf{g}(k) = \sum_{i=1}^7 \Phi_i \mathbf{g}(k-i) + \mathbf{e}_g(k), \quad (C11)$$

where:

$$\Phi_1 = \begin{bmatrix} 0.612 & 0.231 & 0.001 & \cdots & -0.006 & -0.003 \\ 0.035 & 0.609 & -0.135 & \cdots & 0.034 & -0.003 \\ -0.282 & -0.125 & -0.228 & \cdots & 0.060 & -0.010 \\ 0.007 & -0.050 & 0.134 & \cdots & 0.004 & -0.021 \\ -0.006 & 0.139 & 0.232 & \cdots & -0.022 & 0.037 \\ -0.203 & 0.272 & 0.119 & \cdots & -0.004 & -0.003 \\ -0.070 & -0.339 & -0.130 & \cdots & 0.016 & 0.028 \\ 0.224 & 0.471 & 0.002 & \cdots & 0.002 & -0.025 \\ -0.348 & -0.664 & -0.085 & \cdots & 0.204 & -0.018 \\ 0.895 & 0.619 & -0.260 & \cdots & -0.022 & 0.212 \end{bmatrix}, \quad (C12a)$$

$$\Phi_2 = \begin{bmatrix} 0.225 & -0.102 & 0.010 & \cdots & 0.012 & -0.007 \\ -0.067 & -0.056 & 0.024 & \cdots & 0.017 & -0.006 \\ 0.221 & -0.559 & 0.247 & \cdots & 0.020 & -0.000 \\ -0.432 & 0.034 & 0.058 & \cdots & 0.088 & -0.035 \\ 0.447 & 0.084 & -0.069 & \cdots & -0.113 & 0.041 \\ -0.015 & -0.186 & 0.013 & \cdots & 0.017 & -0.001 \\ 0.118 & 0.028 & -0.009 & \cdots & -0.030 & 0.022 \\ -0.218 & -0.090 & 0.051 & \cdots & 0.046 & -0.037 \\ 0.242 & 0.299 & -0.137 & \cdots & 0.161 & 0.023 \\ -0.010 & -0.208 & -0.025 & \cdots & 0.022 & 0.170 \end{bmatrix}, \quad (C12b)$$

$$\Phi_3 = \begin{bmatrix} 0.106 & -0.076 & 0.011 & \cdots & 0.014 & -0.003 \\ -0.116 & 0.050 & 0.100 & \cdots & 0.007 & 0.008 \\ -0.093 & -0.099 & 0.399 & \cdots & 0.073 & 0.000 \\ 0.157 & -0.039 & 0.039 & \cdots & -0.035 & -0.025 \\ -0.133 & 0.044 & -0.169 & \cdots & 0.021 & 0.035 \\ 0.045 & -0.088 & 0.024 & \cdots & 0.003 & 0.005 \\ -0.084 & 0.029 & 0.087 & \cdots & -0.009 & 0.012 \\ -0.013 & -0.127 & -0.006 & \cdots & 0.005 & -0.005 \\ 0.002 & 0.231 & -0.039 & \cdots & 0.122 & 0.004 \\ -0.028 & -0.034 & -0.036 & \cdots & -0.002 & 0.134 \end{bmatrix}, \quad (C12c)$$

$$\Phi_4 = \begin{bmatrix} 0.014 & 0.002 & 0.005 & \cdots & 0.016 & -0.000 \\ -0.025 & 0.090 & 0.011 & \cdots & 0.011 & -0.006 \\ -0.057 & 0.157 & -0.024 & \cdots & 0.076 & -0.008 \\ -0.011 & 0.071 & 0.034 & \cdots & -0.010 & -0.023 \\ 0.025 & -0.141 & -0.011 & \cdots & -0.003 & 0.029 \\ 0.032 & 0.040 & -0.003 & \cdots & 0.006 & 0.011 \\ -0.127 & -0.037 & 0.032 & \cdots & 0.005 & 0.009 \\ 0.029 & 0.002 & 0.016 & \cdots & -0.005 & -0.015 \\ 0.227 & 0.043 & -0.065 & \cdots & 0.064 & 0.006 \\ -0.140 & -0.012 & 0.041 & \cdots & -0.003 & 0.093 \end{bmatrix}, \quad (\text{C12d})$$

$$\Phi_5 = \begin{bmatrix} 0.061 & -0.023 & -0.003 & \cdots & 0.008 & 0.003 \\ -0.003 & 0.123 & 0.034 & \cdots & 0.011 & -0.005 \\ 0.093 & 0.214 & 0.059 & \cdots & 0.086 & -0.009 \\ -0.056 & 0.079 & 0.050 & \cdots & -0.032 & -0.004 \\ 0.051 & -0.150 & -0.077 & \cdots & 0.009 & 0.011 \\ 0.060 & -0.001 & 0.008 & \cdots & 0.005 & 0.004 \\ -0.013 & 0.114 & 0.004 & \cdots & 0.012 & -0.013 \\ -0.085 & -0.104 & 0.043 & \cdots & -0.012 & -0.001 \\ -0.133 & 0.043 & -0.014 & \cdots & 0.030 & 0.015 \\ 0.031 & -0.098 & -0.046 & \cdots & -0.039 & 0.105 \end{bmatrix}, \quad (\text{C12e})$$

$$\Phi_6 = \begin{bmatrix} -0.005 & -0.022 & -0.002 & \cdots & 0.010 & 0.000 \\ -0.009 & 0.058 & -0.002 & \cdots & 0.014 & -0.008 \\ -0.019 & 0.154 & -0.015 & \cdots & 0.082 & -0.002 \\ 0.114 & 0.018 & 0.000 & \cdots & -0.017 & 0.001 \\ -0.108 & -0.046 & -0.001 & \cdots & -0.004 & 0.000 \\ 0.037 & -0.043 & -0.015 & \cdots & 0.005 & -0.003 \\ 0.058 & 0.082 & -0.006 & \cdots & 0.004 & -0.010 \\ -0.025 & -0.065 & 0.005 & \cdots & -0.007 & 0.012 \\ 0.033 & 0.035 & 0.004 & \cdots & -0.002 & -0.026 \\ -0.230 & -0.104 & 0.056 & \cdots & 0.004 & 0.085 \end{bmatrix}, \quad (\text{C12f})$$

$$\Phi_7 = \begin{bmatrix} -0.023 & -0.005 & 0.019 & \cdots & 0.005 & 0.006 \\ 0.172 & 0.114 & -0.035 & \cdots & -0.007 & -0.003 \\ 0.060 & 0.165 & 0.079 & \cdots & 0.037 & 0.025 \\ 0.245 & -0.073 & -0.031 & \cdots & -0.023 & 0.002 \\ -0.279 & 0.067 & -0.005 & \cdots & 0.014 & -0.009 \\ 0.039 & 0.020 & -0.014 & \cdots & 0.002 & 0.001 \\ 0.128 & 0.124 & -0.029 & \cdots & 0.001 & -0.008 \\ 0.083 & -0.099 & -0.037 & \cdots & -0.008 & 0.004 \\ -0.051 & -0.009 & -0.026 & \cdots & -0.011 & 0.003 \\ -0.527 & -0.181 & 0.239 & \cdots & -0.016 & 0.052 \end{bmatrix}. \quad (\text{C12g})$$

References

Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49(268), 765-769.

- Apley, D. W., & Shi, J. (1999). The GLRT for statistical process control of autocorrelated processes. *IIE Transactions*, 31(12), 1123-1134.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: Forecasting and control* (5th ed.). Hoboken, New Jersey: John Wiley & Sons.
- Cai, L., Tian, X., & Chen, S. (2017). Monitoring nonlinear and non-Gaussian processes using Gaussian mixture model-based weighted kernel independent component analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 28(1), 122-135.
- Chan, N. H., & Wei, C. Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics*, 16(1), 367-401.
- Chen, Q., Kruger, U., & Leung, A. Y. T. (2009). Cointegration testing method for monitoring nonstationary processes. *Industrial & Engineering Chemistry Research*, 48(7), 3533-3543.
- Chen, Q., Wynne, R., Goulding, P., & Sandoz, D. (2000). The application of principal component analysis and kernel density estimation to enhance process monitoring. *Control Engineering Practice*, 8(5), 531-543.
- Chigira, H. (2008). A test of cointegration rank based on principal component analysis. *Economics Letters*, 15(9), 693-696.
- Chiremsel, Z., Said, R. N., & Chiremsel, R. (2016). Probabilistic fault diagnosis of safety instrumented systems based on fault tree analysis and Bayesian network. *Journal of Failure Analysis and Prevention*, 16(5), 747-760.
- Davies, B., Jackson, L., & Dunnett, S. (2017). Expert diagnosis of polymer electrolyte fuel cells. *International Journal of Hydrogen Energy*, 42(16), 11724-11734.
- Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for auto-regressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 49(4), 1057-1072.
- Ding, S. X. (2013). *Model-based fault diagnosis techniques: Design schemes, algorithms and tools* (2nd ed.). London, UK: Springer-Verlag.
- Dong, Y., & Qin, S. J. (2018). A novel dynamic PCA algorithm for dynamic data modeling and process monitoring. *Journal of Process Control*, 67(SI), 1-11.
- Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 55(2), 251-276.
- Escribano, A., & Peña, D. (1994). Cointegration and common factors. *Journal of Time Series Analysis*, 15(6), 577-586.

- Fan, W., Cai, G., Zhu, Z. K., Shen, C., Huang, W., & Shang, L. (2015). Sparse representation of transients in wavelet basis and its application in gearbox fault feature extraction. *Mechanical Systems and Signal Processing*, 56-57, 230-245.
- Fan, Y., & Gencay, R. (2010). Unit root tests with wavelets. *Econometric Theory*, 26(5), 1305-1331.
- Fan, J., & Wang, Y. (2014). Fault detection and diagnosis of non-linear non-Gaussian dynamic processes using kernel dynamic independent component analysis. *Information Sciences*, 259, 369-379.
- Gao, Z., Cecati, C., & Ding, S. X. (2015). A survey of fault diagnosis and fault-tolerant techniques – Part II: Fault diagnosis with knowledge-based and hybrid/active approaches. *IEEE Transactions on Industrial Electronics*, 62(6), 3768-3774.
- Ge, Z., & Song, Z. (2013). *Multivariate statistical process control: Process monitoring methods and applications*. London, UK: Springer-Verlag.
- Ge, Z., Song, Z., & Gao, F. (2013). Review of recent research on data-based process monitoring. *Industrial & Engineering Chemistry Research*, 52(10), 3543-3562.
- Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, 13(1), 27-35.
- Hannan, E. J. (1980). The estimation of the order of an ARMA process. *The Annals of Statistics*, 8, 1071-1081.
- Huang, J., & Yan, X. (2015). Dynamic process fault detection and diagnosis based on dynamic principal component analysis, dynamic independent component analysis and Bayesian inference. *Chemometrics and Intelligent Laboratory Systems*, 148, 115-127.
- Jaffel, I., Taouali, O., Harkat, M. F., & Messaoud, H. (2016). Moving window KPCA with reduced complexity for nonlinear dynamic process monitoring. *ISA Transactions*, 64, 184-192.
- Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55(2), 163-172.
- Johansen, S. (1995). *Likelihood-based inference in cointegrated vector auto-regressive models*. Oxford: Oxford University Press.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Kasa, K. (1992). Common stochastic trends in international stock markets. *Journal of Monetary Economics*, 29(1), 95-124.
- Kricheldorf, H. R., Nuyken, O., & Swift, G. (Eds.) (2005). *Handbook of polymer synthesis* (2nd ed.). New York: Marcel Dekker.

- Kruger, U., & Xie, L. (2012). *Statistical monitoring of complex multivariate processes: with applications in industrial process control*. Chichester, UK: John Wiley & Sons Ltd.
- Kruger, U., Zhou, Y., & Irwin, G. W. (2004). Improved principal component monitoring of large-scale processes. *Journal of Process Control*, 14(8), 879-888.
- Ku, W., Storer, R. H., & Georgakis, C. (1995). Disturbance rejection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 30(1), 179-196.
- Lee, J., Qin, S. J., & Lee, I. (2006). Fault detection and diagnosis based on modified independent component analysis. *AIChE Journal*, 52(10), 3501-3504.
- Lee, J. M., Yoo, C., & Lee, I. B. (2004). Statistical process monitoring of dynamic processes based on independent component analysis. *Chemical Engineering Science*, 14(7), 2995-3006.
- Lei, Y., Lin, J., He, Z., & Zuo, M. J. (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 35(1), 108-126.
- Li, H., Gao, Y., Shi, P., & Lam, H. (2016). Observer-based fault detection for nonlinear systems with sensor fault and limited communication capacity. *IEEE Transactions on Automatic Control*, 61(9), 2745-2751.
- Li, G., Qin, S. J., & Zhou, D. (2010). Geometric properties of partial least squares for process monitoring. *Automatica*, 46(1), 204-210.
- Li, G., Qin, S. J., & Zhou, D. (2014). A new method of dynamic latent-variable modeling for process monitoring. *IEEE Transactions on Industrial Electronics*, 61(11), 6438-6445.
- Li, N., & Yang, Y. (2015). Ensemble kernel principal component analysis for improved nonlinear process monitoring. *Industrial & Engineering Chemistry Research*, 54(1), 318-329.
- Lieftucht, D., Kruger, U., Xie, L., Littler, T., Chen, Q., & Wang, S. Q. (2006). Statistical monitoring of dynamic multivariate processes – Part 2. Identifying fault magnitude and signature. *Industrial & Engineering Chemistry Research*, 45(5), 1677-1688.
- Lin, Y., Kruger, U., & Chen, Q. (2017). Monitoring nonstationary dynamic systems using cointegration and common-trends analysis. *Industrial & Engineering Chemistry Research*, 56(31), 8895-8905.
- Liu, J., Luo, W., Yang, X., & Wu, L. (2016). Robust model-based fault diagnosis for PEM fuel cell air-feed system. *IEEE Transactions on Industrial Electronics*, 63(5), 3261-3270.
- MacKinnon, J. G. (1991). Critical values for cointegration tests. In R.F. Engle and C.W.J. Granger (Eds.), *Long-run economic relationships*, pp. 267-276. Oxford: Oxford University Press.
- McFarlane, R. C., Reineman, R. C., Bartee, J. F. & Georgakis, C. (1993). Dynamic simulator for a model IV fluid catalytic cracking unit. *Computers & Chemical Engineering*, 17(3), 275-300.
- Mohammadi, E., & Montazeri-Gh, M. (2015). A fuzzy-based gas turbine fault detection and identification system for full and part-load performance deterioration. *Aerospace Science and Technology*, 46, 82-93.

- Pfaff, B. (2008). *Analysis of integrated and cointegrated time series with R*. New York: Springer Science & Business Media.
- Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-346.
- Qin, S. J. (2012). Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, 36(2), 220-234.
- Rato, T., Reis, M., Schmitt, E., Hubert, M., & De Ketelaere, B. (2016). A systematic comparison of PCA-based statistical process monitoring methods for high-dimensional, time-dependent processes. *AIChE Journal*, 62(5), 1478-1493.
- Severson, K., Chaiwatanodom, P., & Braatz, R. D. (2016). Perspectives on process monitoring of industrial systems. *Annual Reviews in control*, 42, 190-200.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611.
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. UK: Chapman & Hall.
- Stock, J. H., & Watson, M. W. (1988). Testing for common trends. *Journal of the American Statistical Association*, 83(404), 1097-1107.
- Sun, H., Zhang, S., Zhao, C., & Gao, F. (2017). A sparse reconstruction strategy for online fault diagnosis in nonstationary processes with no a priori fault information. *Industrial & Engineering Chemistry Research*, 56(24), 6993-7008.
- Superville, C. R., & Adams, B. M. (1994). An evaluation of forecast-based quality control schemes. *Communications in Statistics – Simulations and Computations*, 23(3), 645-661.
- Tanaka, K. (1996). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. New York: Wiley.
- Wang, X., Kruger, U., & Irwin, G. W. (2005). Process monitoring approach using fast moving window PCA. *Industrial & Engineering Chemistry Research*, 44(15), 5691-5702.
- Wang, X., Kruger, U., & Lennox, B. (2003). Recursive partial least squares algorithms for monitoring complex industrial processes. *Control Engineering Practice*, 11(6), 613-632.
- Wu, C., Guo, C., Xie, Z., Ni, F., & Liu, H. (2018). A signal-based fault detection and tolerance control method of current sensor for PMSM drive. *IEEE Transactions on Industrial Electronics*, 65(12), 9646-9657.
- Xie, L., Kruger, U., Lieftucht, D., Littler, T., Chen, Q., & Wang, S. Q. (2006). Statistical monitoring of dynamic multivariate processes – Part 1. Modeling autocorrelation and cross-correlation. *Industrial & Engineering Chemistry Research*, 45(5), 1659-1676.

- Xie, L., Li, Z., Zeng, J., & Kruger, U. (2016). Block adaptive kernel principal component analysis for nonlinear process monitoring. *AIChE Journal*, 62(12), 4334-4345.
- Yan, R., Gao, R. X., & Chen, X. (2014). Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal processing*, 96(5), 1-15.
- Yin, S., Li, X., Gao, H., & Kaynak, O. (2015). Data-based techniques focused on modern industry: An overview. *IEEE Transactions on Industrial Electronics*, 62(1), 657-667.
- Yu, H., Khan, F., & Garaniya, V. (2016). A sparse PCA for nonlinear fault diagnosis and robust feature discovery of industrial processes. *AIChE Journal*, 62(5), 1494-1513.
- Zhang, Y., Sun, R., & Fan, Y. (2015). Fault diagnosis of nonlinear process based on KCPLS reconstruction. *Chemometrics and Intelligent Laboratory Systems*, 140, 49-60.
- Zhao, Z., Yang, Y., Ding, S. X., & Li, L. (2018). Robust fault detection for observer-based feedback control systems. *Circuits Systems and Signal Processing*, 37(8), 3364-3382.
- Zhong, M., Xue, T., & Ding, S. X. (2018). A survey on model-based fault diagnosis for linear discrete time-varying systems. *Neurocomputing*, 306(51), 51-60.

List of Figures

- Figure 1: Flowchart of the proposed approach for nonstationary and dynamic process monitoring; the new components added to the common trend framework are marked by the shaded boxes. 39
- Figure 2: Section of reference data for FCCU application. 40
- Figure 3: Estimating the number of stationary factors for FCCU application. 40
- Figure 4: Monitoring charts for fault condition of FCCU application: (a) T_{ns}^2 statistic based on the compensated VARI model residuals of nonstationary factors, (b) T_s^2 statistic based on the compensated VAR model residuals of stationary factors, (c) Hotelling's T^2 statistic based on the PCA model, and (d) SPE statistic based on the PCA model. 41
- Figure 5: Monitoring charts for the normal operating data of FCCU application: (a) T_{ns}^2 statistic based on the compensated VARI model residuals of nonstationary factors, (b) T_s^2 statistic based on the compensated VAR model residuals of stationary factors, (c) Hotelling's T^2 statistic based on the PCA model, and (d) SPE statistic based on the PCA model. 42
- Figure 6: Section of reference data for application to polymerization process. 42
- Figure 7: Section of recorded data for application to polymerization process describing a fault condition. 43
- Figure 8: Monitoring charts for fault condition of polymerization process: (a) T_{ns}^2 statistic based on the compensated VARI model residuals of nonstationary factors, (b) T_s^2 statistic based on the compensated VAR model residuals of stationary factors, (c) Hotelling's T^2 statistic based on the PCA model, and (d) SPE statistic based on the PCA. 43
- Figure 9: Monitoring charts for fault condition of polymerization process: (a) T_{ns}^2 statistic based on VARI model residuals of nonstationary factors without application of compensation scheme, and (b) T_s^2 statistic based on VAR model residuals of stationary factors without application of compensation scheme. 44
- Figure 10: Section of reference data for application to distillation unit. 44
- Figure 11: Monitoring charts for fault condition of distillation process (a) T_{ns}^2 statistic based on the compensated VARI model residuals of nonstationary factors, (b) T_s^2 statistic based on the compensated VAR model residuals of stationary factors, (c) Hotelling's T^2 statistic based on the PCA model, and (d) SPE statistic based on the PCA model. 45
- Figure 12: Monitoring charts for the normal operating data of distillation unit (a) T_{ns}^2 statistic based on the VARI model residuals of nonstationary factors, and (b) T_s^2 statistic based on the VAR model residuals of stationary factors. The control limits were computed from a scaled F-distribution. 46

Figures

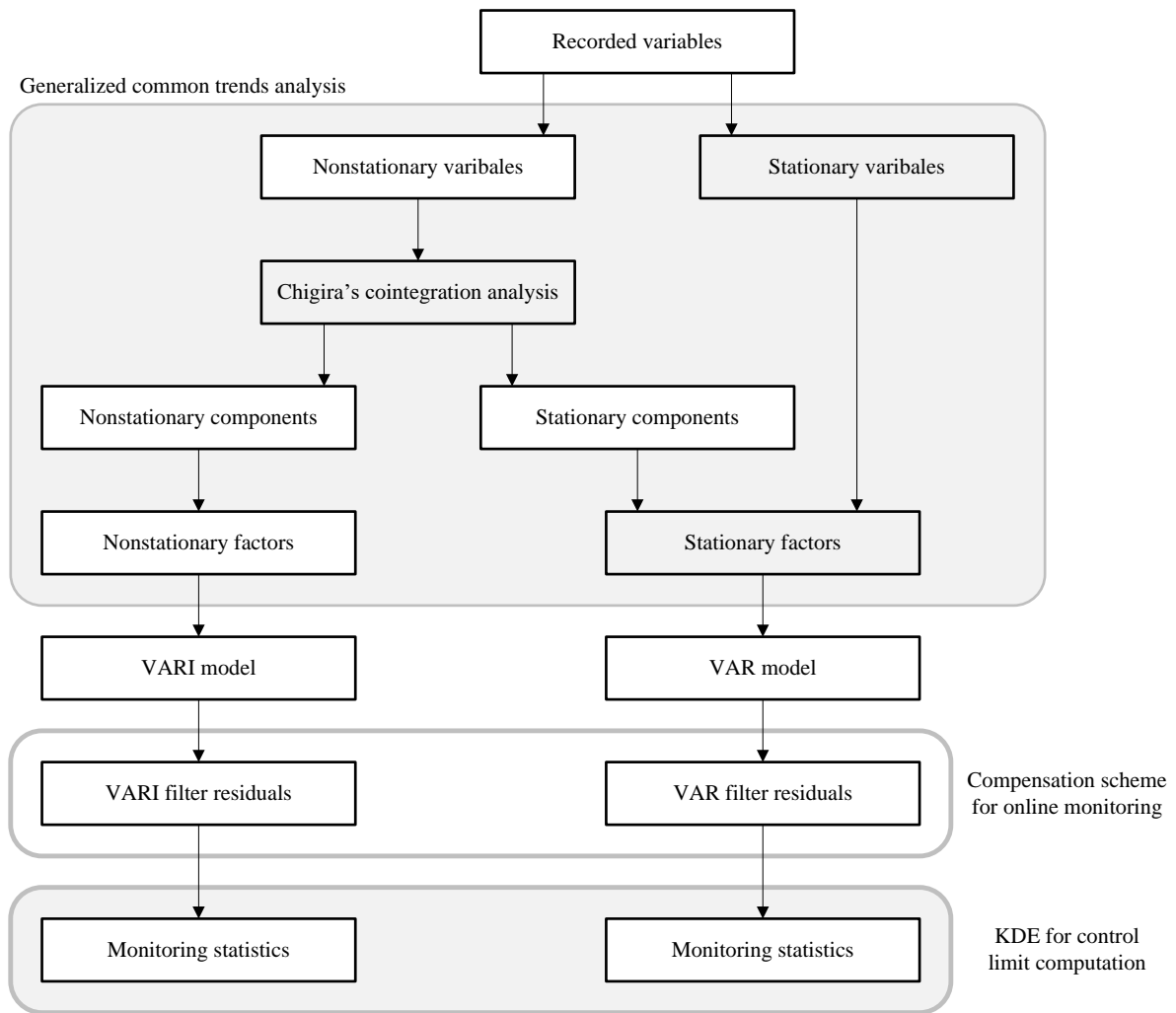


Figure 1: Flowchart of the proposed approach for nonstationary and dynamic process monitoring; the new components added to the common trend framework are marked by the shaded boxes.

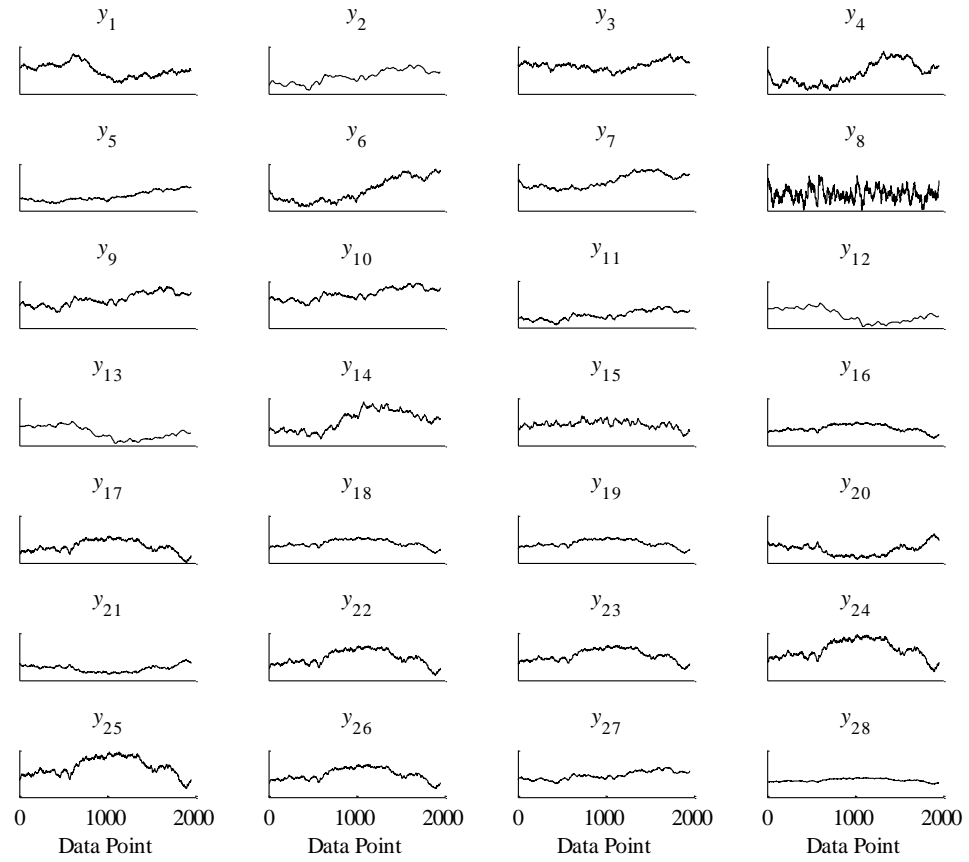


Figure 2: Section of reference data for FCCU application.

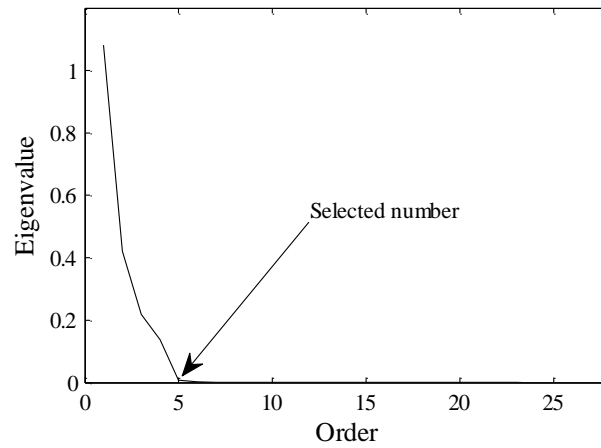


Figure 3: Estimating the number of stationary factors for FCCU application.

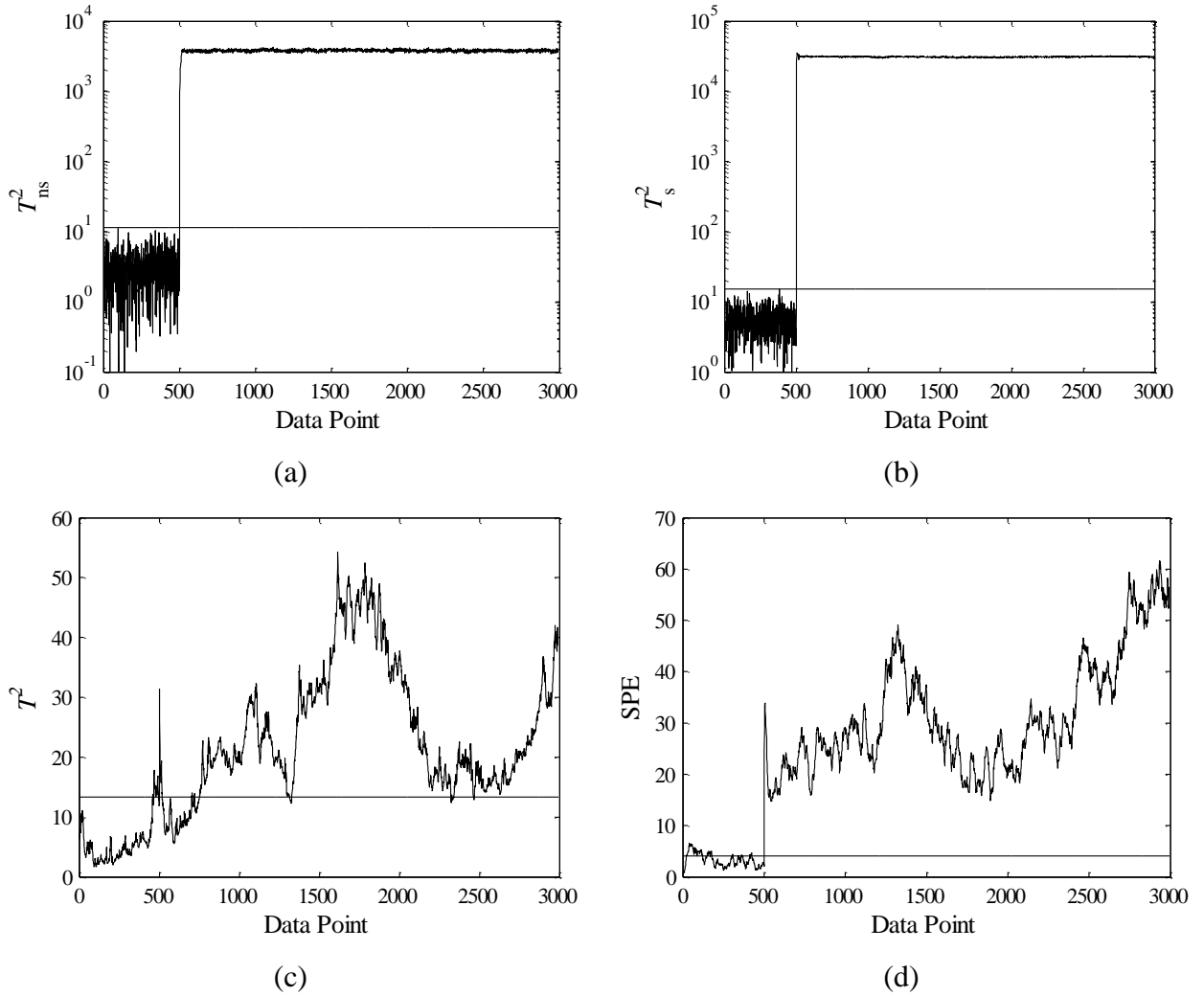


Figure 4: Monitoring charts for fault condition of FCCU application: (a) T^2_{ns} statistic based on the compensated VARI model residuals of nonstationary factors, (b) T^2_s statistic based on the compensated VAR model residuals of stationary factors, (c) Hotelling's T^2 statistic based on the PCA model, and (d) SPE statistic based on the PCA model.

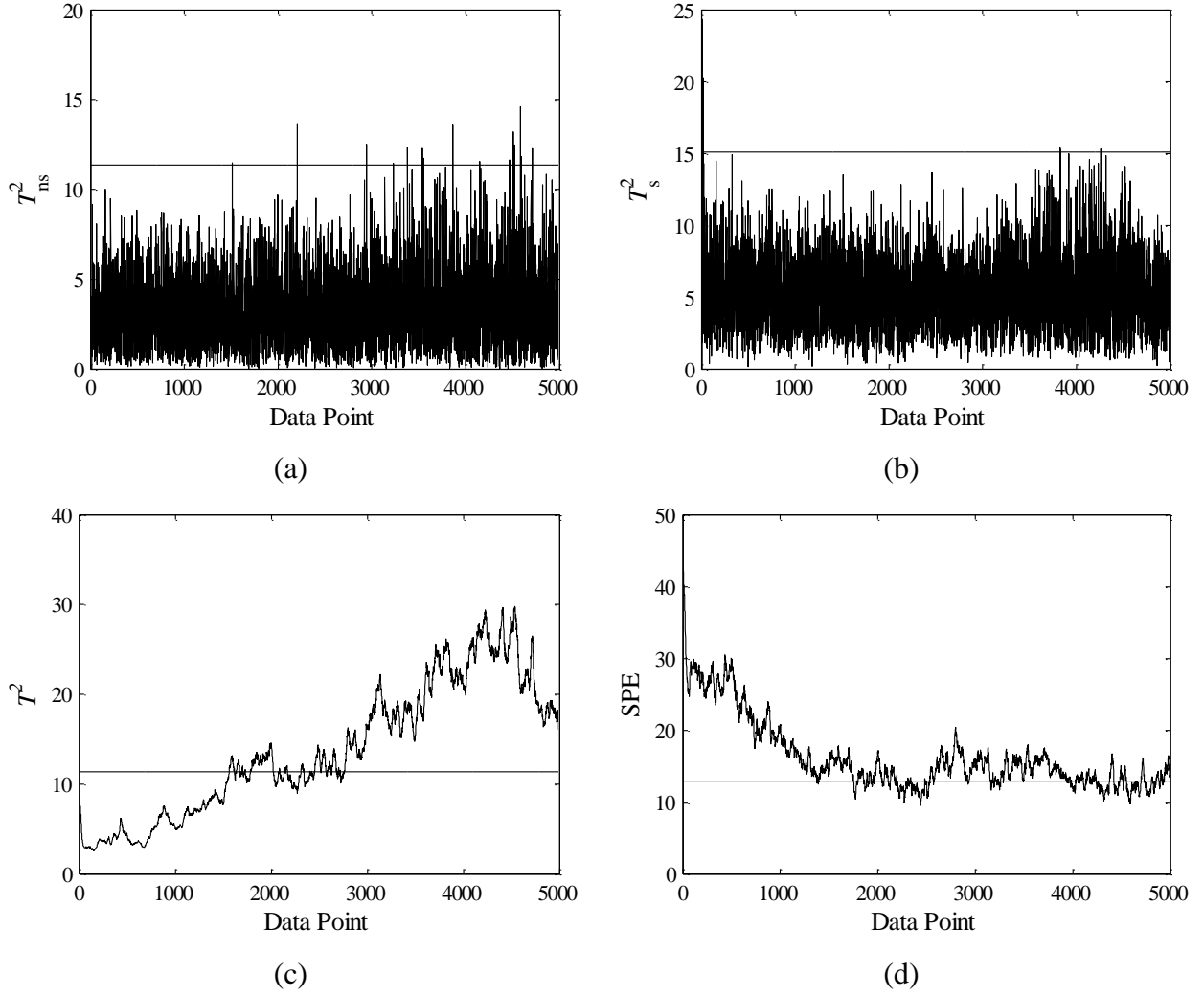


Figure 5: Monitoring charts for the normal operating data of FCCU application: (a) T^2_{ns} statistic based on the compensated VARI model residuals of nonstationary factors, (b) T^2_s statistic based on the compensated VAR model residuals of stationary factors, (c) Hotelling's T^2 statistic based on the PCA model, and (d) SPE statistic based on the PCA model.

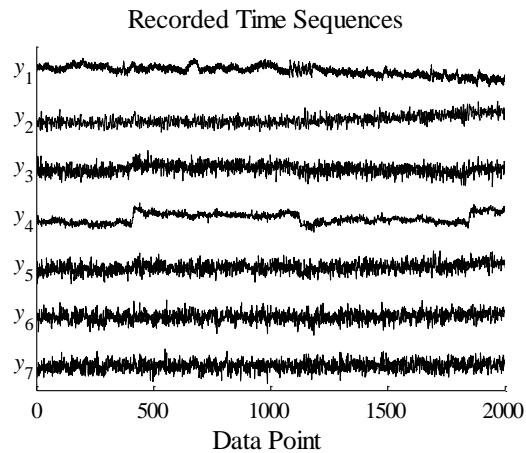


Figure 6: Section of reference data for application to polymerization process.

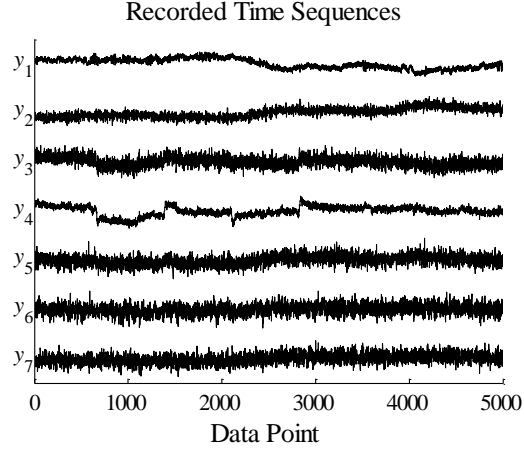


Figure 7: Section of recorded data for application to polymerization process describing a fault condition.

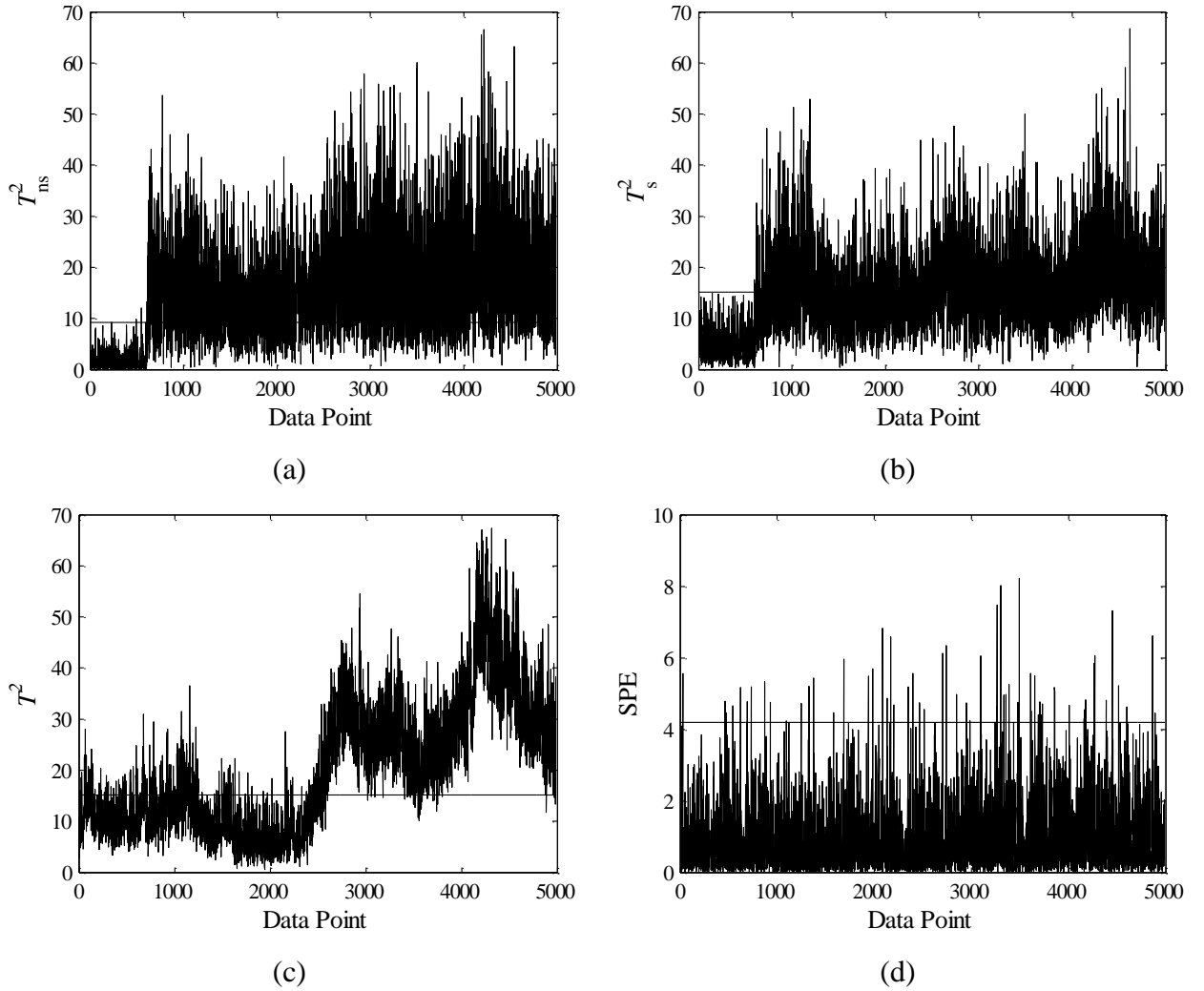


Figure 8: Monitoring charts for fault condition of polymerization process: (a) T^2_{ns} statistic based on the compensated VARI model residuals of nonstationary factors, (b) T^2_s statistic based on the compensated VAR model residuals of stationary factors, (c) Hotelling's T^2 statistic based on the PCA model, and (d) SPE statistic based on the PCA.

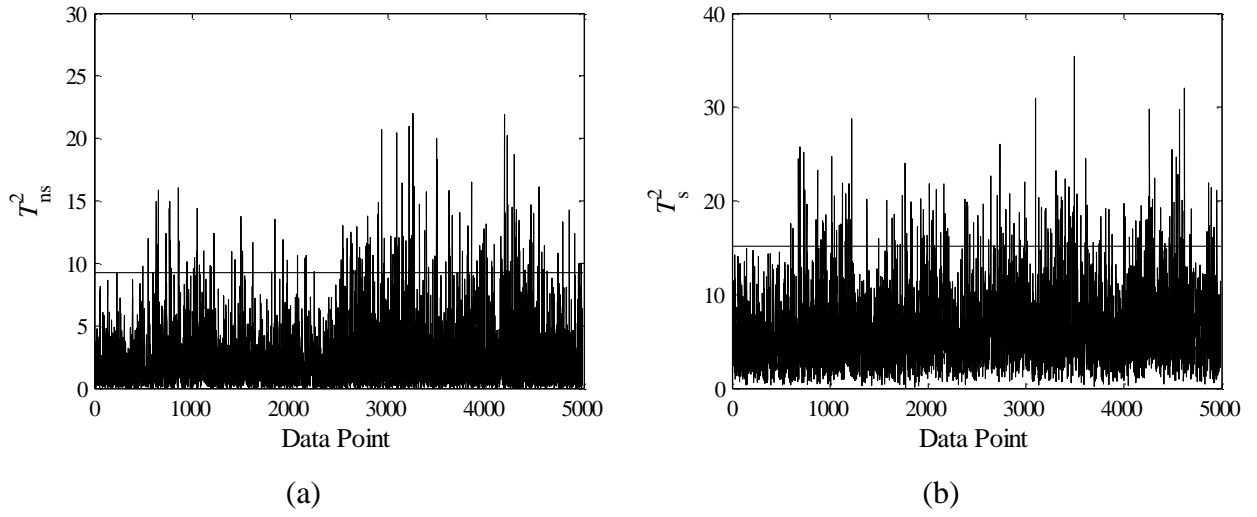


Figure 9: Monitoring charts for fault condition of polymerization process: (a) T^2_{ns} statistic based on VARI model residuals of nonstationary factors without application of compensation scheme, and (b) T^2_s statistic based on VAR model residuals of stationary factors without application of compensation scheme.

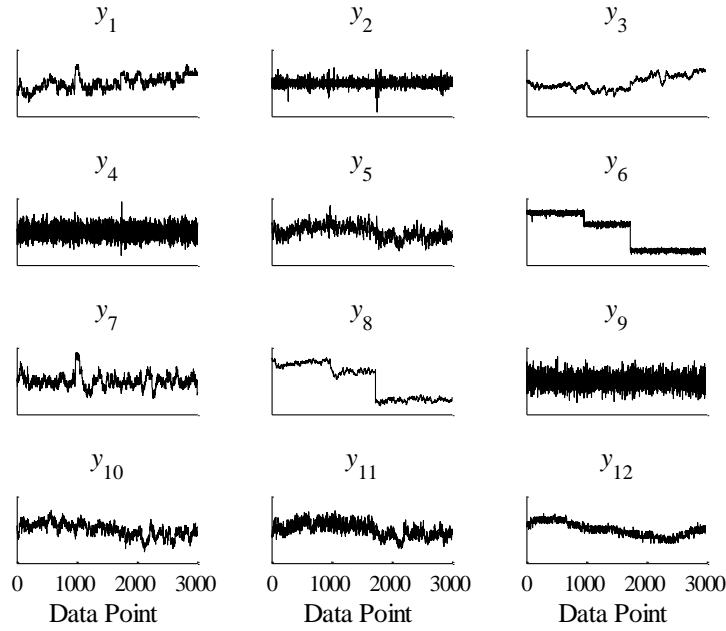


Figure 10: Section of reference data for application to distillation unit.

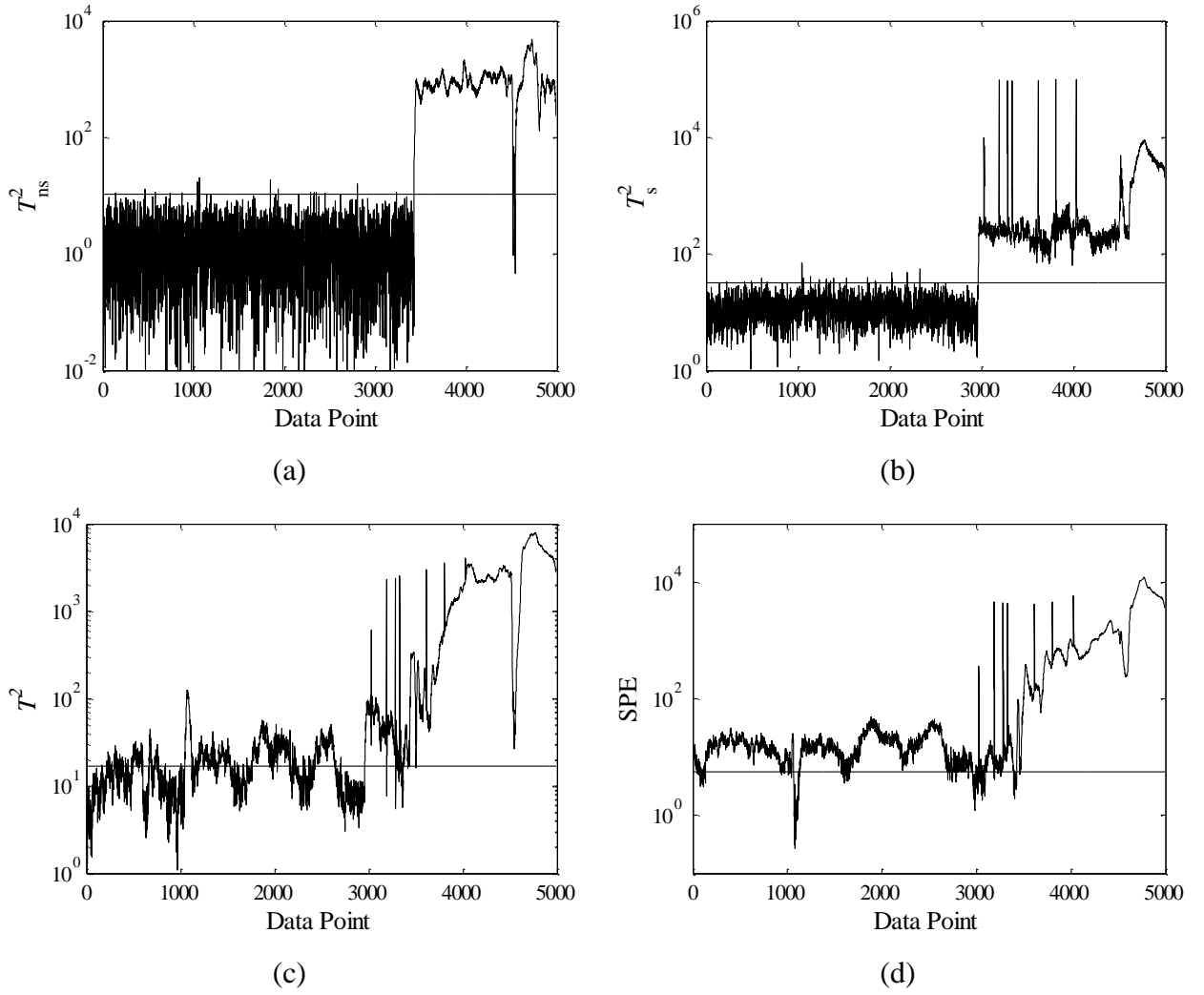


Figure 11: Monitoring charts for fault condition of distillation process (a) T^2_{ns} statistic based on the compensated VARI model residuals of nonstationary factors, (b) T^2_s statistic based on the compensated VAR model residuals of stationary factors, (c) Hotelling's T^2 statistic based on the PCA model, and (d) SPE statistic based on the PCA model.

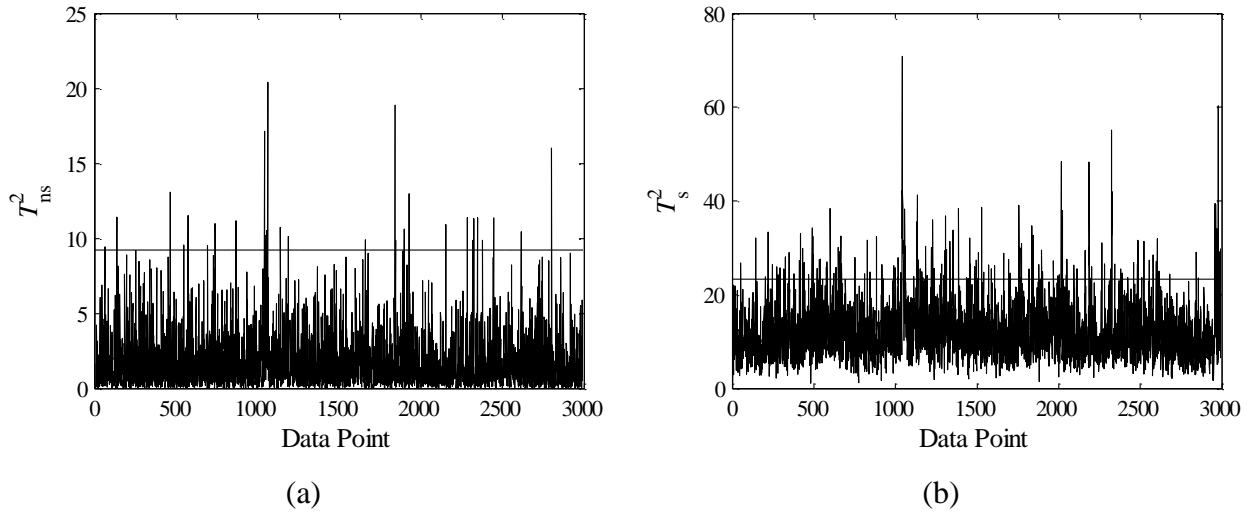


Figure 12: Monitoring charts for the normal operating data of distillation unit (a) T^2_{ns} statistic based on the VARI model residuals of nonstationary factors, and (b) T^2_s statistic based on the VAR model residuals of stationary factors. The control limits were computed from a scaled F -distribution.

List of Tables

Table 1: Recorded variable set of FCCU simulator	48
Table 2: Recorded variable set of spinning process	49
Table 3: Recorded variable set of distillation unit	49
Table 4: Results of ADF test applied to recorded variable set of FCCU simulator	49
Table 5: Results of ADF test applied to recorded variables set of polymerization process	52
Table 6: Results of ADF test applied to recorded variables set of distillation unit	53
Table 7: Results of AD test applied to residuals of VARI and VAR models of distillation unit	53
Table 8: Summary of monitoring results for industrial application studies	54

Tables

Table 1: Recorded variable set of FCCU simulator

variable number	description
1	wash oil feed flowrate y_1
2	total fresh feed flowrate y_2
3	slurry flowrate y_3
4	furnace fuel flowrate y_4
5	preheat train outlet temperature y_5
6	fresh feed temperature to reactor y_6
7	furnace firebox temperature y_7
8	riser temperature y_8
9	wet gas compressor suction pressure y_9
10	wet gas compressor suction flowrate y_{10}
11	wet gas flowrate to vapor recovery unit y_{11}
12	regenerator bed temperature y_{12}
13	stack gas (cyclone exit) temperature y_{13}
14	stack gas O ₂ concentration y_{14}
15	stack gas CO concentration y_{15}
16	standpipe catalyst level y_{16}
17	combustion air blower suction flowrate y_{17}
18	combustion air blower throughput y_{18}
19	combustion air flowrate y_{19}
20	combustion air blower suction pressure y_{20}
21	combustion air blower discharge pressure y_{21}
22	lift air blower suction flowrate y_{22}
23	lift air blower speed y_{23}
24	lift air blower throughput y_{24}
25	lift air flowrate y_{25}
26	lift air blower discharge pressure y_{26}
27	wet gas compressor suction valve y_{27}
28	regenerator stack gas valve y_{28}

Table 2: Recorded variable set of spinning process

variable number	description (controller output)
1	gas flow y_1
2	positional heater gas temperature y_2
3	plenum gas temperature y_3
4	head metal temperature y_4
5	cell wall temperature y_5
6	upper mid temperature y_6
7	lower mid temperature y_7

Table 3: Recorded variable set of distillation unit

variable number	description
1	tray 14 temperature y_1
2	column overhead pressure y_2
3	tray 2 temperature y_3
4	reflux vessel level y_4
5	butane product flow y_5
6	reflux flow y_6
7	fresh feed temperature y_7
8	butane product flow y_8
9	reboiler vessel level y_9
10	reboiler steam flow y_{10}
11	fresh feed flow y_{11}
12	reboiler outlet temperature y_{12}

Table 4: Results of ADF test applied to recorded variable set of FCCU simulator

variable	AR order	t statistic	critical value	test result
wash oil feed flowrate y_1	3	-1.2346	-3.4335	nonstationary
wash oil feed flowrate (differ.) Δy_1	2	-22.7284	-3.4335	stationary

total fresh feed flowrate y_2	4	-1.7450	-3.4335	nonstation- ary
total fresh feed flowrate (differ.) Δy_2	3	-6.1039	-3.4335	stationary
slurry flowrate y_3	3	-2.0412	-3.4335	nonstation- ary
slurry flowrate (differ.) Δy_3	2	-22.2858	-3.4335	stationary
furnace fuel flowrate y_4	3	-0.9473	-3.4335	nonstation- ary
furnace fuel flowrate (differ.) Δy_4	2	-21.8103	-3.4335	stationary
preheat train outlet temperature y_5	4	0.1269	-3.4335	nonstation- ary
preheat train outlet temperature (differ.) Δy_5	3	-21.6231	-3.4335	stationary
fresh feed temperature to reactor y_6	4	0.1145	-3.4335	nonstation- ary
fresh feed temperature to reactor (differ.) Δy_6	3	-19.9465	-3.4335	stationary
furnace firebox temperature y_7	3	-0.6527	-3.4335	nonstation- ary
furnace firebox temperature (differ.) Δy_7	2	-21.5982	-3.4335	stationary
riser temperature y_8	3	-6.0901	-3.4335	stationary
wet gas compressor suction pressure y_9	3	-1.0770	-3.4335	nonstation- ary
wet gas compressor suction pressure (differ.) Δy_9	2	-22.0375	-3.4335	stationary
wet gas compressor suction flowrate y_{10}	3	-1.0792	-3.4335	nonstation- ary
wet gas compressor suction flowrate (differ.) Δy_{10}	2	-22.0268	-3.4335	stationary
wet gas flowrate to vapor recovery unit y_{11}	3	-1.0773	-3.4335	nonstation- ary
wet gas flowrate to vapor recovery unit (differ.) Δy_{11}	2	-22.0286	-3.4335	stationary
regenerator bed temperature y_{12}	4	-1.7319	-3.4335	nonstation- ary
regenerator bed temperature (differ.) Δy_{12}	3	-6.3228	-3.4335	stationary

stack gas (cyclone exit) temperature y_{13}	5	-1.5365	-3.4335	nonstation- ary
stack gas (cyclone exit) temperature (differ.) Δy_{13}	4	-7.1144	-3.4335	stationary
stack gas O ₂ concentration y_{14}	5	-1.3042	-3.4335	nonstation- ary
stack gas O ₂ concentration (differ.) Δy_{14}	4	-14.1361	-3.4335	stationary
stack gas CO concentration y_{15}	5	-3.2802	-3.4335	nonstation- ary
stack gas CO concentration (differ.) Δy_{15}	4	-9.1446	-3.4335	stationary
standpipe catalyst level y_{16}	3	-1.1643	-3.4335	nonstation- ary
standpipe catalyst level (differ.) Δy_{16}	2	-25.2470	-3.4335	stationary
combustion air blower suction flowrate y_{17}	3	-1.3136	-3.4335	nonstation- ary
combustion air blower suction flowrate (differ.) Δy_{17}	2	-21.9990	-3.4335	stationary
combustion air blower throughput y_{18}	3	-1.3136	-3.4335	nonstation- ary
combustion air blower throughput (differ.) Δy_{18}	2	-21.9989	-3.4335	stationary
combustion air flowrate y_{19}	3	-1.3135	-3.4335	nonstation- ary
combustion air flowrate (differ.) Δy_{19}	2	-22.0024	-3.4335	stationary
combustion air blower suction pressure y_{20}	3	-1.3137	-3.4335	nonstation- ary
combustion air blower suction pressure (differ.) Δy_{20}	2	-22.0010	-3.4335	stationary
combustion air blower discharge pressure y_{21}	3	-1.3148	-3.4335	nonstation- ary
combustion air blower discharge pressure (differ.) Δy_{21}	2	-22.0084	-3.4335	stationary
lift air blower suction flowrate y_{22}	4	-1.2256	-3.4335	nonstation- ary
lift air blower suction flowrate (differ.) Δy_{22}	3	-20.3767	-3.4335	stationary
lift air blower speed y_{23}	4	-1.2227	-3.4335	nonstation- ary

lift air blower speed (differ.) Δy_{23}	3	-20.4341	-3.4335	stationary
lift air blower throughput y_{24}	4	-1.2256	-3.4335	nonstationary
lift air blower throughput (differ.) Δy_{24}	3	-20.3767	-3.4335	stationary
lift air flowrate y_{25}	4	-1.2256	-3.4335	nonstationary
lift air flowrate (differ.) Δy_{25}	3	-20.3784	-3.4335	stationary
lift air blower discharge pressure y_{26}	4	-1.2215	-3.4335	nonstationary
lift air blower discharge pressure (differ.) Δy_{26}	3	-20.4633	-3.4335	stationary
wet gas compressor suction valve y_{27}	3	-1.0676	-3.4335	nonstationary
wet gas compressor suction valve (differ.) Δy_{27}	2	-22.0840	-3.4335	stationary
regenerator stack gas valve y_{28}	3	-1.3066	-3.4335	nonstationary
regenerator stack gas valve (differ.) Δy_{28}	3	-20.4084	-3.4335	stationary

Table 5: Results of ADF test applied to recorded variables set of polymerization process

variable	AR order	t statistic	critical value	test result
gas flow y_1	12	-2.6918	-3.4334	nonstationary
gas flow (differ.) Δy_1	11	-18.6111	-3.4334	stationary
positional heater gas temperature y_2	13	-3.1408	-3.4334	nonstationary

positional heater gas temperature (differ.) Δy_2	12	-23.0953	-3.4334	stationary
plenum gas temperature y_3	10	-5.9883	-3.4334	stationary
head metal temperature y_4	6	-2.7970	-3.4334	nonstationary
head metal temperature (differ.) Δy_4	5	-27.8522	-3.4334	stationary
cell wall temperature y_5	10	-7.6864	-3.4334	stationary
upper mid temperature y_6	2	-21.9061	-3.4334	stationary
lower mid temperature y_7	2	-22.5715	-3.4334	stationary

Table 6: Results of ADF test applied to recorded variables set of distillation unit

variable	AR order	t statistic	critical value	test result
tray 14 temperature y_1	5	-4.4290	-3.4323	stationary
column overhead pressure y_2	2	-37.5481	-3.4323	stationary
tray 2 temperature y_3	4	-1.0703	-3.4323	nonstationary
tray 2 temperature (differ.) Δy_3	3	-36.7893	-3.4323	stationary
reflux vessel level y_4	5	-19.7659	-3.4323	stationary
butane product flow y_5	4	-7.0230	-3.4323	stationary
reflux flow y_6	5	-0.9983	-3.4323	nonstationary
reflux flow (differ.) Δy_6	4	-35.8335	-3.4323	stationary
fresh feed temperature y_7	5	-5.4104	-3.4323	stationary
butane product flow y_8	1	-1.4268	-3.4323	nonstationary
butane product flow (differ.) Δy_8	0	-58.4056	-3.4323	stationary
reboiler vessel level y_9	9	-18.3997	-3.4323	stationary
reboiler steam flow y_{10}	3	-5.0873	-3.4323	stationary
fresh feed flow y_{11}	5	-5.1321	-3.4323	stationary
reboiler outlet temperature y_{12}	17	-1.9700	-3.4323	nonstationary
reboiler outlet temperature (differ.) Δy_{12}	16	-20.4594	-3.4323	stationary

Table 7: Results of AD test applied to residuals of VARI and VAR models of distillation unit

variable	statistic	critical value	test result
e_{f1}	1.2612	1.0338	non-normal
e_{f2}	0.2940	1.0338	normal

e_{g1}	0.3720	1.0338	normal
e_{g2}	0.6278	1.0338	normal
e_{g3}	1.6915	1.0338	non-normal
e_{g4}	0.7235	1.0338	normal
e_{g5}	0.5559	1.0338	normal
e_{g6}	0.2299	1.0338	normal
e_{g7}	0.9293	1.0338	normal
e_{g8}	1.1159	1.0338	non-normal
e_{g9}	1.2709	1.0338	non-normal
e_{g10}	1.4040	1.0338	non-normal

Table 8: Summary of monitoring results for industrial application studies

case	approach	statistic	occurrence / detection	fault data		normal data	introduced improvements		
			time index	FAR	MAR	FAR	data structure	Chigira test	KDE
FCCU	proposed	T_{ns}^2	501 / 501	0%	0%	0.30%	√	√	×
		T_s^2	501 / 501	0%	0%	0.12%			
	PCA	T^2	501 / 501	6.00%	0%	57.56%	N/A		
		SPE	501 / 501	29.60%	0%	76.50%			
polymerization	proposed	T_{ns}^2	601 / 628	0.20%	17.77%	N/A	√	√	×
		T_s^2	601 / 616	0.20%	28.52%				
	PCA	T^2	601 / 618	14.17%	38.20%	N/A	N/A		
		SPE	601 / 620	0.67%	98.64%				
	uncompensated	T_{ns}^2	N/A	N/A	98.20%	N/A	√	√	×
		T_s^2			95.57%				
distillation	proposed	T_{ns}^2	3001 / 3424	0.53%	21.75%	N/A	√	√	√
		T_s^2	3001 / 3001	0.68%	0%				
	PCA	T^2	3001 / 3001	44.00%	2.20%	N/A	N/A		
		SPE	3001 / 3001	93.30%	5.50%				
	without KDE	T_{ns}^2	N/A	0.93%	N/A	N/A	√	√	×
		T_s^2		5.21%					